GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2001, 23:48:51 ; Search time 2434.93 Seconds
(without alignments)
3620.891 Million cell updates/sec

Title: US-09-558-474-1
Perfect score: 570
Sequence: 1 ATGCTGGGGAGCAGAGCTGT.....CAGCAACCCTGAGTCCCTAA 570

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	570	100.0	570	10	AX097621	AX097621 Sequence
	2	570	100.0	1025	89	AF301620	AF301620 Homo sapi
	3	566.8	99.4	1026	85	AB030000	AB030000 Homo sapi
	4	566.8	99.4	1055	9	AX048200	AX048200 Sequence
	5	391	68.6	1203	10	AX097623	AX097623 Sequence
	6	391	68.6	1358	94	AF301619	AF301619 Mus muscu
	7	192.8	33.8	4039	85	AB030001	AB030001 Homo sapi
c	8	192.8	33.8	155929	69	AC025574	AC025574 Homo sapi

	9	149.2	26.2	2300	7	AB030002 AB030002 Sus scrof
c	10	149.2	26.2	4649	7	AB004061 AB004061 domestic
	11	40.8	7.2	11696	3	SCPSCL X54107 S.clavulige
c	12	40.6	7.1	937	53	CNS06SJ1 AL413315 T3 end of
	13	40	7.0	101270	92	HS355C18 AL022327 Human DNA
	14	39.4	6.9	233231	67	AC022220 AC022220 Homo sapi
c	15	39	6.8	55815	62	AC011100 AC011100 Homo sapi
c	16	39	6.8	129091	64	AC016033 AC016033 Homo sapi
	17	39	6.8	178022	65	AC018348 AC018348 Homo sapi
	18	38.8	6.8	110000	84	LMFLCHR18_02 Continuation (3 of
	19	38.6	6.8	2279	89	AF217995 AF217995 Homo sapi
	20	38.4	6.7	85127	69	AC025978 AC025978 Homo sapi
c	21	38.4	6.7	191496	61	AC009130 AC009130 Homo sapi
c	22	38.2	6.7	173695	79	AL354898 AL354898 Homo sapi
c	23	38	6.7	121162	9	AX039602 AX039602 Sequence
	24	38	6.7	153766	78	AL139424 AL139424 Homo sapi
	25	38	6.7	176150	82	AP001098 AP001098 Homo sapi
c	26	38	6.7	179119	80	AL358013 AL358013 Homo sapi
	27	37.6	6.6	25025	65	AC020357 AC020357 Drosophil
	28	37.6	6.6	160346	4	AC010578 AC010578 Drosophil
	29	37.6	6.6	160903	4	AC007580 AC007580 Drosophil
	30	37.6	6.6	258778	5	AE003813 AE003813 Drosophil
c	31	37.2	6.5	172706	81	AL391135 AL391135 Homo sapi
c	32	37.2	6.5	179483	87	AC019230 AC019230 Homo sapi
	33	37	6.5	57335	63	AC015272 AC015272 Drosophil
c	34	37	6.5	64620	76	AC083794 AC083794 Homo sapi
	35	37	6.5	188459	4	AC008312 AC008312 Drosophil
	36	37	6.5	197597	4	AC011253 AC011253 Drosophil
	37	37	6.5	309023	5	AE003671 AE003671 Drosophil
	38	36.8	6.5	108653	93	HSDJ863C7 AL049761 Human DNA
	39	36.8	6.5	160536	71	AC040953 AC040953 Homo sapi
c	40	36.6	6.4	159299	71	AC027825 AC027825 Homo sapi
	41	36.6	6.4	233735	61	AC010443 AC010443 Homo sapi
c	42	36.4	6.4	1934	14	PHGCR1 X04335 Petunia grp
c	43	36.4	6.4	221437	88	AC034242 AC034242 Homo sapi
c	44	36.2	6.4	84912	89	AL137840 AL137840 Human DNA
	45	36.2	6.4	204480	71	AC034265 AC034265 Mus muscu

ALIGNMENTS

```
RESULT   1
AX097621
LOCUS       AX097621      570 bp    DNA              PAT       30-MAR-2001
DEFINITION  Sequence 1 from Patent WO0118051.
ACCESSION   AX097621
VERSION     AX097621.1  GI:13514261
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Oppmann,B., de waal Malefyt,R., Rennick,D.M., Kastelein,R.A.,
            Wiekowski,M.T., Lira,S.A. and Narula,S.K.
  TITLE     Mammalian interleukin-12 p40 and interleukin b30. Combinations
            thereof. Antibodies. Uses in pharmaceutical compositions
  JOURNAL   Patent: WO 0118051-A 1 15-MAR-2001;
            SCHERING CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="surmised Homo sapiens"
     CDS             1. .570
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35429.1"
                     /db_xref="GI:13514262"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVAQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
```

```
     mat_peptide     64. .567
BASE COUNT      121 a    168 c    157 g    124 t
ORIGIN
```

Query Match 100.0%; Score 570; DB 10; Length 570;
Best Local Similarity 100.0%; Pred. No. 1.5e-152;
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240

Qy     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300

Qy     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360

Qy     361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420

Qy     421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480

Qy     481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540

Qy     541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           |||||||||||||||||||||||||||||||
Db     541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
```

```
RESULT   2
AF301620
LOCUS       AF301620     1025 bp    mRNA             PRI       04-DEC-2000
DEFINITION  Homo sapiens interleukin 23 p19 subunit mRNA, complete cds.
ACCESSION   AF301620
VERSION     AF301620.1  GI:11528342
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Oppmann,B., Lesley,R., Blom,B., Timans,J.C., Xu,Y., Hunte,B.,
            Vega,F., Yu,N., Wang,J., Singh,K., Zonin,F., Vaisberg,E.,
            Churakova,T., Liu,M., Gorman,D., Wagner,J., Zurawski,S., Liu,Y.,
            Abrams,J.S., Moore,K.W., Rennick,D., de Waal-Malefyt,R., Hannum,C.,
            Bazan,J.F. and Kastelein,R.A.
  TITLE     Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with
            biological activities similar as well as distinct from IL-12
  JOURNAL   Immunity 13 (5), 715-725 (2000)
  MEDLINE   20567322
REFERENCE   2  (bases 1 to 1025)
  AUTHORS   Kastelein,R.A., Gorman,D., Timans,J.C., Oppmann,B. and Bazan,J.F.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) Molecular Biology, DNAX Research Institute,
            901 California Avenue, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             151. .720
                     /note="combines with the p40 subunit of IL-12 to form a
                     novel, biologically active, composite cytokine, IL-23"
                     /codon_start=1
                     /product="interleukin 23 p19 subunit"
                     /protein_id="AAG37232.1"
                     /db_xref="GI:11528343"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVAQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
BASE COUNT      262 a    252 c    274 g    237 t
ORIGIN
```

```
  Query Match             100.0%;  Score 570;  DB 89;  Length 1025;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-152;
  Matches  570;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 210

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 270

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 330

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     331 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 390

Qy     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     391 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 450

Qy     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 510

Qy     361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     511 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 570

Qy     421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     571 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 630

Qy     481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     631 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 690

Qy     541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||
Db     691 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 720
```

```
RESULT   3
AB030000
LOCUS       AB030000     1026 bp    mRNA            PRI       04-APR-2000
DEFINITION  Homo sapiens mRNA for SGRF, complete cds.
ACCESSION   AB030000
VERSION     AB030000.1  GI:7416072
KEYWORDS    SGRF.
SOURCE      Homo sapiens fetal spleen cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1026)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 1026)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .1026
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="fetal"
                     /tissue_type="spleen"
     gene            144. .713
                     /gene="SGRF"
     CDS             144. .713
                     /gene="SGRF"
                     /codon_start=1
                     /product="SGRF"
                     /protein_id="BAA93686.1"
                     /db_xref="GI:7416073"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVGQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
     sig_peptide     144. .200
                     /gene="SGRF"
     mat_peptide     201. .710
                     /gene="SGRF"
                     /product="SGRF"
     polyA_signal    996. .1001
BASE COUNT      268 a    249 c    272 g    237 t
ORIGIN
```

```
  Query Match             99.4%;  Score 566.8;  DB 85;  Length 1026;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-151;
  Matches  568;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     144 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 203

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 263

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     264 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 323

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 383

Qy     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 443

Qy     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     444 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 503
```

```
Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504 GTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 563

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    564 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 623

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    624 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 683

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    684 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 713


RESULT   4
AX048200
LOCUS       AX048200     1055 bp    DNA             PAT       15-DEC-2000
DEFINITION  Sequence 40 from Patent WO0070049.
ACCESSION   AX048200
VERSION     AX048200.1  GI:11876990
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1055)
  AUTHORS   Tang,Y.T., Yue,H., Lal,P., Burford,N., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Lu,D.A. and Patterson,C.
  TITLE     Extracellular signaling molecules
  JOURNAL   Patent: WO 0070049-A 40 23-NOV-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1055
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 2933038CB1"
BASE COUNT      284 a    254 c    278 g    239 t
ORIGIN


  Query Match             99.4%;  Score 566.8;  DB 9;  Length 1055;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-151;
  Matches  568;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 230

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 290

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 350

Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 410

Qy    241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 470

Qy    301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 530

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 GTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 590

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    591 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 650

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    651 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 710

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    711 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 740


RESULT   5
AX097623
LOCUS       AX097623     1203 bp    DNA             PAT       30-MAR-2001
DEFINITION  Sequence 3 from Patent WO0118051.
ACCESSION   AX097623
VERSION     AX097623.1  GI:13514264
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Oppmann,B., de waal Malefyt,R., Rennick,D.M., Kastelein,R.A.,
            Wiekowski,M.T., Lira,S.A. and Narula,S.K.
  TITLE     Mammalian interleukin-12 p40 and interleukin b30. Combinations
            thereof. Antibodies. Uses in pharmaceutical compositions
  JOURNAL   Patent: WO 0118051-A 3 15-MAR-2001;
            SCHERING CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="surmised Mus sp."
     CDS             113. .703
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35430.1"
                     /db_xref="GI:13514265"
                     /translation="MLDCRAVIMLWLLPWVTQGLAVPRSSSPDWAQCQQLSRNLCMLA
                     WNAHAPAGHMNLLREEEDEETKNNVPRIQCEDGCDPQGLKDNSQFCLQRIRQGLAFYK
                     HLLDSDIFKGEPALLPDSPMEQLHTSLLGLSQLLQPEDHPRETQQMPSLSSSQQWQRP
                     LLRSKILRSLQAFLAIAARVFAHGAATLTEPLVPTA"
     mat_peptide     176. .700
BASE COUNT      336 a    310 c    312 g    245 t
ORIGIN


  Query Match             68.6%;  Score 391;  DB 10;  Length 1203;
  Best Local Similarity   81.8%;  Pred. No. 3e-101;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          |||||||   |||||||  |||||  ||||  | ||||  ||||||||    ||||||||
Db    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||| || | |||||  ||||  |||| |||||||||||||||| || | ||| |||
Db    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
          ||||  ||  |||||||  ||||||  ||| ||  ||||||  |||||||||||
Db    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292

Qy    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
          | ||||||||||||| |||| |||| |||| ||||||||||| ||||||  ||||||||
Db    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352

Qy    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
          |||||||||| |||||||| ||||||||||||||||||||||| ||| ||||||  ||||
```

```
Db    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412

Qy    298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||||
Db    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472

Qy    358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417
          ||  ||| ||| |||||  |||||||||| || ||||||||||||| ||||| |||| |
Db    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532

Qy    418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477
          |||| | ||||||| || ||||| || ||||| ||| | || |||| ||||||| | |
Db    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592

Qy    478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537
          |||||||| | ||| |||||||| |||||||||||||| |||| |||| |||||||||
Db    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652

Qy    538 TTTGCCCATGGAGCAGCAACCCTGAGT 564
          |||||||| |||||||||| |||| |
Db    653 TTTGCCCACGGAGCAGCAACTCTGACT 679


RESULT   6
AF301619
LOCUS       AF301619     1358 bp    mRNA            ROD       04-DEC-2000
DEFINITION  Mus musculus interleukin 23 p19 subunit mRNA, complete cds.
ACCESSION   AF301619
VERSION     AF301619.1  GI:11528340
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1358)
  AUTHORS   Oppmann,B., Lesley,R., Blom,B., Timans,J.C., Xu,Y., Hunte,B.,
            Vega,F., Yu,N., Wang,J., Singh,K., Zonin,F., Vaisberg,E.,
            Churakova,T., Liu,M., Gorman,D., Wagner,J., Zurawski,S., Liu,Y.,
            Abrams,J.S., Moore,K.W., Rennick,D., de Waal-Malefyt,R., Hannum,C.,
            Bazan,J.F. and Kastelein,R.A.
  TITLE     Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with
            biological activities similar as well as distinct from IL-12
  JOURNAL   Immunity 13 (5), 715-725 (2000)
  MEDLINE   20567322
REFERENCE   2  (bases 1 to 1358)
  AUTHORS   Kastelein,R.A., Gorman,D., Timans,J.C., Oppmann,B. and Bazan,J.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) Molecular Biology, DNAX Research Institute,
            901 California Avenue, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .1358
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             113. .703
                     /note="four helical bundle cytokine; combines with the p40
                     subunit of IL-12 to form a novel, biologically active,
                     composite cytokine, IL-23"
                     /codon_start=1
                     /product="interleukin 23 p19 subunit"
                     /protein_id="AAG37231.1"
                     /db_xref="GI:11528341"
                     /translation="MLDCRAVIMLWLLPWVTQGLAVPRSSSPDWAQCQQLSRNLCMLA
                     WNAHAPAGHMNLLREEEDEETKNNVPRIQCEDGCDPQGLKDNSQFCLQRIRQGLAFYK
                     HLLDSDIFKGEPALLPDSPMEQLHTSLLGLSQLLQPEDHPRETQQMPSLSSSQQWQRP
                     LLRSKILRSLQAFLAIAARVFAHGAATLTEPLVPTA"
BASE COUNT      383 a    326 c    338 g    311 t
ORIGIN


  Query Match           68.6%;  Score 391;  DB 94;  Length 1358;
  Best Local Similarity  81.8%;  Pred. No. 3e-101;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          |||||||   ||||||| |||||  |||| | |||| ||||||||     ||||||||
Db    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          |||||||| || | ||||| |||| |||| ||||||||||||||||| || | ||| |||
Db    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
          ||||  || |||||||  |||||| ||| || |||||| ||| ||   ||||||||||
Db    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292

Qy    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
          |  |||||||||||| |||| |||| |||| |||||||||| |||||| ||||||||
Db    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352

Qy    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
          |||||||||| ||||||||| |||||||||||||||||||||| ||| ||||||  ||||
Db    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412

Qy    298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||||
Db    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472

Qy    358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417
          ||  ||| ||| |||||  |||||||||| || ||||||||||||| ||||| |||| |
Db    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532

Qy    418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477
          |||| | ||||||| || ||||| || ||||| ||| | || |||| ||||||| | |
Db    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592

Qy    478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537
          |||||||| | ||| |||||||| |||||||||||||| |||| |||| |||||||||
Db    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652

Qy    538 TTTGCCCATGGAGCAGCAACCCTGAGT 564
          |||||||| |||||||||| |||| |
Db    653 TTTGCCCACGGAGCAGCAACTCTGACT 679


RESULT   7
AB030001
LOCUS       AB030001     4039 bp    DNA             PRI       04-APR-2000
DEFINITION  Homo sapiens gene for SGRF, complete cds.
ACCESSION   AB030001
VERSION     AB030001.1  GI:7416074
KEYWORDS    SGRF.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 4039)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q13"
     TATA_signal     38. .43
     exon            318. .622
                     /product="SGRF"
```

```
     gene            join(461. .622,842. .940,1107. .1253,1359. .1520)
                     /gene="SGRF"
     CDS             join(461. .622,842. .940,1107. .1253,1359. .1520)
                     /gene="SGRF"
                     /codon_start=1
                     /product="SGRF"
                     /protein_id="BAA93687.1"
                     /db_xref="GI:7416075"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVGQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
     exon            842. .940
                     /gene="SGRF"
                     /product="SGRF"
     exon            1107. .1253
                     /gene="SGRF"
                     /product="SGRF"
     exon            1359. .1826
                     /product="SGRF"
     polyA_signal    1803. .1808
BASE COUNT     1098 a    940 c   1029 g    972 t
ORIGIN


  Query Match             33.8%;  Score 192.8;  DB 85;  Length 4039;
  Best Local Similarity   74.3%;  Pred. No. 1.9e-44;
  Matches  309;  Conservative    0;  Mismatches    2;  Indels  105;  Gaps    1;

Qy     260 AGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGG 319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 AGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGG 1164

Qy     320 ATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCT 379
           ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    1165 ATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCT 1224

Qy     380 CCCTACTGGGCCTCAGCCAACTCCTG---------------------------------- 405
           |||||||||||||||||||||||||
Db    1225 CCCTACTGGGCCTCAGCCAACTCCTGCAGGTATGAAGTAGGGGCGTGGAGGATGGGGGCT 1284

Qy     406 ---------------------------------------------------------- 405

Db    1285 TGCAGGTGTCAGAGACAGAGGGTTGGGGGTTAAGGGTTTAGAGTCTTCTCTGACTGTGTC 1344

Qy     406 -----------CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 454
                       ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 CTATGTCCTTTCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 1404

Qy     455 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 514
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 1464

Qy     515 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 1520


RESULT   8
AC025574/c
LOCUS       AC025574   155929 bp    DNA             HTG       06-SEP-2000
DEFINITION  Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC025574
VERSION     AC025574.10  GI:9438393
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155929)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155929)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 25, 2000 this sequence version replaced gi:8571540.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HALV
                Center clone name: RP11-348M3
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 6% of reads
                Chemistry: Dye-terminator Big Dye: 94% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 133622 bases at least Q40
                Consensus quality: 145733 bases at least Q30
                Consensus quality: 150334 bases at least Q20
                Estimated insert size: 150694; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    36972: contig of 36972 bp in length
            *    36973    37072: gap of unknown length
            *    37073    64901: contig of 27829 bp in length
            *    64902    65001: gap of unknown length
            *    65002    91014: contig of 26013 bp in length
            *    91015    91114: gap of unknown length
            *    91115   105122: contig of 14008 bp in length
            *   105123   105222: gap of unknown length
            *   105223   113963: contig of 8741 bp in length
            *   113964   114063: gap of unknown length
            *   114064   124894: contig of 10831 bp in length
            *   124895   124994: gap of unknown length
            *   124995   132612: contig of 7618 bp in length
            *   132613   132712: gap of unknown length
            *   132713   139659: contig of 6947 bp in length
            *   139660   139759: gap of unknown length
            *   139760   143764: contig of 4005 bp in length
```

```
            *   143765   143864: gap of unknown length
            *   143865   148062: contig of 4198 bp in length
            *   148063   148162: gap of unknown length
            *   148163   151000: contig of 2838 bp in length
            *   151001   151100: gap of unknown length
            *   151101   153374: contig of 2274 bp in length
            *   153375   153474: gap of unknown length
            *   153475   154810: contig of 1336 bp in length
            *   154811   154910: gap of unknown length
            *   154911   155929: contig of 1019 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-348M3"
BASE COUNT    40876 a  36387 c  36205 g  41119 t   1342 others
ORIGIN


  Query Match             33.8%;  Score 192.8;  DB 69;  Length 155929;
  Best Local Similarity   74.3%;  Pred. No. 1.6e-44;
  Matches  309;  Conservative    0;  Mismatches    2;  Indels  105;  Gaps    1;

Qy     260 AGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGG 319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63431 AGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGG 63372

Qy     320 ATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCT 379
           ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db   63371 ATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCT 63312

Qy     380 CCCTACTGGGCCTCAGCCAACTCCTG---------------------------------- 405
           |||||||||||||||||||||||||
Db   63311 CCCTACTGGGCCTCAGCCAACTCCTGCAGGTATGAAGTAGGGGCGTGGAGGATGGGGGCT 63252

Qy     406 ---------------------------------------------------------- 405

Db   63251 TGCAGGTGTCAGAGACAGAGGGTTGGGGGTTAAGGGTTTAGAGTCTTCTCTGACTGTGTC 63192

Qy     406 -----------CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 454
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63191 CTATGTCCTTTCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 63132

Qy     455 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 514
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63131 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 63072

Qy     515 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63071 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 63016


RESULT   9
AB030002
LOCUS       AB030002     2300 bp    DNA             MAM       04-APR-2000
DEFINITION  Sus scrofa gene for SGRF, complete cds.
ACCESSION   AB030002
VERSION     AB030002.1  GI:7416076
KEYWORDS    SGRF.
SOURCE      Sus scrofa DNA.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2300)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 2300)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .2300
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     TATA_signal     227. .232
     gene            join(433. .603,820. .921,1590. .1736,1837. .1998)
                     /gene="SGRF"
     CDS             join(433. .603,820. .921,1590. .1736,1837. .1998)
                     /gene="SGRF"
                     /codon_start=1
                     /product="SGRF"
                     /protein_id="BAA93688.1"
                     /db_xref="GI:7416077"
                     /translation="MLGSRAVMLMLLLLLLPWTSQGRAVPEGSSPAWAQGQQLSQQLC
                     TLAWTAHLPMGHVDLPREEGDDETTSEVPHIQCGDGCDPQGLRDNSQSCLQRIHQGLV
                     FYEKLLGSDIFTGEPSLHPDGSVGQLHASLLGLRQLLQPEGHHWETEQTPSPSPSQPW
                     QRLLLRLKILRSLQAFVAVAARVFAHGAATLSQ"
     polyA_signal    2265. .2270
BASE COUNT      597 a    563 c    589 g    551 t
ORIGIN


  Query Match             26.2%;  Score 149.2;  DB 7;  Length 2300;
  Best Local Similarity   67.7%;  Pred. No. 6e-32;
  Matches  279;  Conservative    0;  Mismatches   33;  Indels  100;  Gaps    1;

Qy     259 CAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCG 318
           |||| ||||||||||||||||||||| || ||| ||||||||||||||||||| || ||
Db    1587 CAGTCCTGCTTGCAAAGGATCCACCAAGGCCTGGTTTTTTATGAGAAGCTGCTGGGCTCA 1646

Qy     319 GATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCC 378
           || |||||||||||||||||||| | ||||||| || ||||||  |||||||| |||
Db    1647 GACATTTTCACAGGGGAGCCTTCTCTACACCCTGATGGCTCTGTGGGCCAGCTTCACGCC 1706

Qy     379 TCCCTACTGGGCCTCAGCCAACTCCTG---------------------------------- 405
           ||||||||||||||||| |||||| ||
Db    1707 TCCCTACTGGGCCTCAGGCAACTCTTGCAGGTATAAACTAGGGGCCTGGAGCGGGGGGTG 1766

Qy     406 ---------------------------------------------------------- 405

Db    1767 GCTTGCAGGTGTCAGAGACAGTGCCTGGGGGGTGAAGGATCCTCTCTACTAGTGTCCTGT 1826

Qy     406 -------CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAG 458
                   ||||| |||||||||||||||||||||| ||||||  |||||||  |||||||||
Db    1827 GTCCTTTCAGCCCGAGGGTCACCACTGGGAGACTGAGCAGACGCCAAGCCCCAGTCCCAG 1886

Qy     459 CCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGT 518
           |||||| ||||| || |||||||||||| |||| ||||||||||||||||||||||||||
Db    1887 CCAGCCCTGGCAACGCCTCCTTCTCCGCCTCAAGATCCTTCGCAGCCTCCAGGCCTTTGT 1946

Qy     519 GGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||| |||||||||||| |||||||||||||||||| ||||| |  |||
Db    1947 GGCTGTAGCTGCCCGGGTCTTCGCCCATGGAGCAGCAACTCTGAGCCAGTAA 1998


RESULT  10
AB004061/c
LOCUS       AB004061     4649 bp    mRNA            MAM       05-FEB-1999
DEFINITION  domestic pig mRNA for STAT2, complete cds.
ACCESSION   AB004061
VERSION     AB004061.1  GI:2189979
KEYWORDS    STAT2.
SOURCE      Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 4649)
  AUTHORS   Ito,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1997) to the DDBJ/EMBL/GenBank databases.
```

```
•          Yoshiyasu Ito, National Institute of Animal Industry, Animal Genome
           Research Team; 2 Ikenodai, Kukizaki-machi, Inashiki-gun, Ibaragi
           305-0901, Japan (E-mail:yoshiito@niai.affrc.go.jp,
           Tel:81-298-38-8627)
REFERENCE  2  (sites)
▲AUTHORS   Ito,Y., MIKAWA,S., Kobayashi,E., Wada,Y. and MINEZAWA,M.
  TITLE    Pig STAT2 cDNA
  JOURNAL  Unpublished (1997)
FEATURES             Location/Qualifiers
     source          1. .4649
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /cell_type="thoracic aortic cell"
                     /dev_stage="adult"
                     /tissue_type="muscle"
     CDS             82. .2676
                     /codon_start=1
                     /product="STAT2"
                     /protein_id="BAA20332.1"
                     /db_xref="GI:2189980"
                     /translation="MAQWEMLQNLDSPFQDQLHQLYSESLLPVDVRQYLAVWIEDQNW
                     QEAALGNDDSKANMLFFHFLDQLNYDCGRCGQDPECLLLQHNLRKFYRDIQAIPQGPT
                     RLAEMIFNLLLEEKRILIQAQRAQLEQQEPALEAPGENQQHEIESRILELRAMMEKLV
                     KSISQLKDQQDIFCFRYKIKASAKTHSLDPHRTRQEQVLQETLNELDKRRKEVLDASK
                     ALLGRLTTLIELLLPKLEEWKVQQQKACIGAPMDGELEQLEKWFTAEAKLLFHLRQLL
                     KELKGLSSVVKYDEDLLFKGVDLLKAQVTELLQRLLHRAFIVETQPCMPQTPHRPLIL
                     KTGSKFTVRTRLLVRLQEGNESLTAEVSIDRNPPKSQGFRKFNILTSNRKTLTPEKGQ
                     SQGLIWDFGYLTLMEQRSGGAGKGNNKGPLGVTEELHIISFTVKYTYQGLKQELTTDT
                     LPVVIISNMNQLSIAWASILWFNLLSSDPQNQQFFSSPPKAPWNLLGPALSWQFSSHV
                     GQGLNSDQLGMLRDKLFGQNSSTEGLSLSWVDFIKRESPPGKLPFWTWLDKILDLVHD
                     HLKDLWKDGHIMGFVSRSEERRLLKKTISGTFLLRFSETLEGGITCSWVEHQDDDKVL
                     IYSLQPFTKEVLQSLPLTKIISQYQLLTEENIPENPLRFLYPRIPRDEAFGCYNQEKA
                     NLQERKKYLKHKLIVVSNRQVDELQQPPELKLEPDLESLELDLGLAPGPEPGVGLDLE
                     PLLEAGLDLGPELEPMLQSTLEPVLEPILDEVLQGVPEPNLGPELKLEPILEPVSQPV
                     PEPDLPHDLRHLNTDEMQIFRNSMRIEEIMPNGDPLLASQNTNVDEACVFHRSHFYTD
                     GPLIPSDY"
     polyA_site      3622
                     /note="Alternative Poly adenilation site"
     polyA_site      4649
                     /note="21 a nucleotides"
BASE COUNT     1234 a   1157 c   1194 g   1064 t
ORIGIN


  Query Match             26.2%;  Score 149.2;  DB 7;  Length 4649;
  Best Local Similarity   67.7%;  Pred. No. 5.8e-32;
  Matches  279;  Conservative    0;  Mismatches   33;  Indels  100;  Gaps    1;

Qy     259 CAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCG 318
           |||| |||||||||||||||||||||| || ||| |||||||||||||||||| || ||
Db    4109 CAGTCCTGCTTGCAAAGGATCCACCAAGGCCTGGTTTTTTATGAGAAGCTGCTGGGCTCA 4050

Qy     319 GATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCC 378
           || |||||||||||||||||||||| | ||||||| || ||||||  |||||||| |||
Db    4049 GACATTTTCACAGGGGAGCCTTCTCTACACCCTGATGGCTCTGTGGGCCAGCTTCACGCC 3990

Qy     379 TCCCTACTGGGCCTCAGCCAACTCCTG------------------------------ 405
           |||||||||||||||| |||||| ||
Db    3989 TCCCTACTGGGCCTCAGGCAACTCTTGCAGGTATAAACTAGGGGCCTGGAGCGGGGGGTG 3930

Qy     406 ------------------------------------------------------------ 405

Db    3929 GCTTGCAGGTGTCAGAGACAGTGCCTGGGGGGTGAAGGATCCTCTCTACTAGTGTCCTGT 3870

Qy     406 -------CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAG 458
                |||||  |||||||||||||||||||||| ||||||  ||||||| |||||||||
Db    3869 GTCCTTTCAGCCCGAGGGTCACCACTGGGAGACTGAGCAGACGCCAAGCCCCAGTCCCAG 3810

Qy     459 CCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGT 518
           |||||| ||||| || |||||||||||| |||| ||||||||||||||||||||||||||
Db    3809 CCAGCCCTGGCAACGCCTCCTTCTCCGCCTCAAGATCCTTCGCAGCCTCCAGGCCTTTGT 3750

Qy     519 GGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           |||||||| |||||||||| |||||||||||||||||| ||||| |  |||
Db    3749 GGCTGTAGCTGCCCGGGTCTTCGCCCATGGAGCAGCAACTCTGAGCCAGTAA 3698


RESULT  11
SCPSCL
LOCUS       SCPSCL      11696 bp    DNA             BCT       30-JUN-1993
DEFINITION  S.clavuligerus linear plasmid pSCL (complete sequence).
ACCESSION   X54107 S51598
VERSION     X54107.1  GI:48758
KEYWORDS    inverted repeat; linear plasmid.
SOURCE      Streptomyces clavuligerus.
  ORGANISM  Plasmid Streptomyces clavuligerus
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 11696)
  AUTHORS   Roy,K.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1990) Roy K.L., Dept of Microbiology, Unversity
            of Alberta, Biological Sciences Bldg Rm M330, Edmonton Alberta,
            Canada T6G 2E9
REFERENCE   2  (bases 1 to 11696)
  AUTHORS   Wu,X. and Roy,K.L.
  TITLE     Complete nucleotide sequence of a linear plasmid from Streptomyces
            clavuligerus and characterization of its RNA transcripts
  JOURNAL   J. Bacteriol. 175 (1), 37-52 (1993)
  MEDLINE   93106972
FEATURES             Location/Qualifiers
     source          1. .11696
                     /organism="Streptomyces clavuligerus"
                     /plasmid="pSCL"
                     /strain="NRRL 3585"
                     /db_xref="taxon:1901"
     CDS             complement(1446. .2237)
                     /note="ORF1"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38040.1"
                     /db_xref="GI:48759"
                     /db_xref="SPTREMBL:Q05073"
                     /translation="MAMTRDALLNQKGGVAKTTTTLHLGGTLAAAGRRTLLVDLDGQG
                     NLTTALKLPRLNPAGELTLARAVLEGCDRAGAQSLVRRHSENLWVIPSALDLFTLPRQ
                     LHSARSKEERLSWVLEHLEDGFDHCLVDCRPALDVDTDNVLRWAHDILIPVDVDEFSI
                     EALELLLGQVQTLVRETRMAPPRYRGLVINRVARPFSGFHQAVYDALHRLPLPVVGEI
                     PLRTAVAEAKNQGQTLAQYARTADVARMFRDLAVKAGYLPEAPAA"
     CDS             complement(4062. .5417)
                     /note="ORF2"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38041.1"
                     /db_xref="GI:581632"
                     /db_xref="SPTREMBL:Q05074"
                     /translation="MSMESEELRAAEVREAEDAVLGTLLRPEARDLVPDLVSALAPTD
                     FREPRYGELWDLLASLWARGTPCDPVTVSAVLRKRGDLERIGGAVWLHALHDAVPTTA
                     HALRYAEIVAEESRWRALLAAGVRTVTAAQRREGTAAEAAETAVGAVREVRDRGLAAA
                     DAPPADLIDFLDQADDEPEWVIPGVLARWDRLMITAGEGGGKSLLLRQILVRAAAGLH
                     PWRRARIEPVRVLLVDAENAASQVRPWLRRMAAAAAGEGAEVARGRMHIEVRPEGLDL
                     TTAADRAWLARTVERVRPDLMSIGPLYKLTTSGASDEEAARPLMSALEMVRQASDGAA
                     MLLEAHSPHAAPGVRRRDLRPIGSSLWLRWPEFGFGLSPATGMGAEELRLMDWLPWRG
                     PRSEREWPAQFCEGGTWPWTAVQYAGNVPIPAGTLTEEQAVAAGLIPEQPEFWDDQQI
                     G"
     CDS             complement(6349. .6786)
                     /note="ORF5"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38042.1"
                     /db_xref="GI:48761"
                     /db_xref="SPTREMBL:Q05076"
                     /translation="MEAALHLRFAQLRQHGEWFDFGSDNPVLKVEEAITAGLPEIPAQ
                     QPSHRVIPPRPSVSKDARCRCGHELSFHRGTAGCTVAGWDEWLDCSCERFRAAEADWH
                     VGMPACRHCRHSAQDHLGIEGPCSGKGACGCYHFSASGVGVAR"
     prim_transcript complement(6931. .7914)
                     /note="large RNA transcript"
```

```
     CDS             complement(7139. .7885)
                     /note="ORFL"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38043.1"
                     /db_xref="GI:48762"
                     /db_xref="SPTREMBL:Q05077"
                     /translation="MPKIERPLPPYMQVIAHIRDQIDSGELRAGDLIPSDRNIAQEWG
                     ISRATAQKVITALKAQGLVEAVQGSGTRVRAAADHLHRSGRDRAASVRRTGRIYTAGE
                     YARIVSAELAPAPADVAEVLSVQEGAAVIRRIRITYNAENQPISASTSWYDGNLAAAA
                     PKLLKTERIKEGSWAYLEEQTGMTAVYGQDRISARMATEQDAQLLGIELPAAIKEART
                     ILRDANSVAVEYGVSINGNGRESIYDYDVS"
     promoter        complement(7921. .7951)
                     /note="promoter for large RNA transcript"
     prim_transcript complement(7951. .8400)
                     /note="small RNA transcript"
     CDS             8114. .9367
                     /note="ORF3"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38044.1"
                     /db_xref="GI:581633"
                     /db_xref="SPTREMBL:Q05075"
                     /translation="MLEAPRPLSDPTCLARRSDVSIVAERRVVAQVGVRPVVAGLAEQ
                     VEEWEREDDAQGPTDGPEGGVRRWSERQRELGARRAVAVPGQRRVTPSYTAPDRVDLR
                     TAEDVPLTPVLAARVAALVEEHGDHLVRYVAARLRDAEWAYWARAEDVAQDVWLDVAR
                     GRVPELLEEVPDLPWPRLAAAAKWSLLDNTRTRRRREWLVRVPEDRSADEVLEALAGP
                     GPDSTVCAVDELMEPEPEPGGWAPACYAERIAALPPRQREVLELRCTEGMTTPAIAAR
                     LGISRQSVDRALRFAVTALGSPGTVVRQRSRGAGQPLPAGWERVLDRLPNRTQRDVVR
                     LRAGGASFGELGEQLGLHRGYAHELYTRALRSLREMVQDQRLDPVPAAPARSKPKPHP
                     APAGCARTSCVSGCYLRTARTGAAS"
     promoter        complement(8411. .8442)
                     /note="promoter for small RNA transcript"
     CDS             9696. .10679
                     /note="ORF4"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA38045.1"
                     /db_xref="GI:48764"
                     /db_xref="SPTREMBL:Q05083"
                     /translation="MGREETGPGVRPQWVLRTLLGGAVAVGAVAMAASAVTLAELGHA
                     VGWVQWDGRLSWSLPVAVDLLALVAGAVWLSPGMTARSRSLARGITLGAVAGSVALNA
                     IGHLVESGDMKVTPLLRIAVSAVPPLVAALVVHLVGVVLAPVPGADHGQDQAVELEEQ
                     REAATISAAHDPADGTTADRHDGTTARNAPAPQPRPAPAPDAPVRVPAAPAADARTAA
                     DDTPAPPPVPAAAQTGPPRVTARLGDRPDDSQIWERPAPVRDWDRPDLPADCAPGTEP
                     LLLTPDQSRARILWALQNTDWTQKRIGEFAGRSASVVNTRKAEAIRDGILI"
BASE COUNT     1590 a    4235 c    4179 g    1692 t
ORIGIN


  Query Match             7.2%;  Score 40.8;  DB 3;  Length 11696;
  Best Local Similarity   51.7%;  Pred. No. 0.66;
  Matches   93;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy     336 GCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAG 395
           || | | | || |||  |  | |    |  |   |  ||   | || ||| ||| |||
Db    5222 GCGTCCCCCGCGCCCACAGCGACGCCAGCAGATCCCACAACTCCCCGTACCGGGGCTCCC 5281

Qy     396 CCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCC 455
           || |||  | |    ||||     ||  |      || ||||| || |||||| | ||
Db    5282 GGAAATCCGTCGGGGCCAGGGCGGAGACCAGATCCGGCACCAGATCCCGAGCCTCCGGCC 5341

Qy     456 CAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTT 515
             ||||     ||||  | | | ||| || ||     |||  || |||  |||  | |
Db    5342 GCAGCAGCGTCCCCAGCACCGCGTCCTCGGCCTCCCGCACCTCAGCCGCCCGCAGCTCCT 5401


RESULT 12
CNS06SJ1/c
LOCUS       CNS06SJ1      937 bp    DNA             STS       10-JAN-2001
DEFINITION  T3 end of clone AW0AA023F10 of library AW0AA from strain CLIB 89 of
            Yarrowia lipolytica, sequence tagged site.
ACCESSION   AL413315
VERSION     AL413315.1  GI:12184937
KEYWORDS    STS.
SOURCE      Yarrowia lipolytica.
  ORGANISM  Yarrowia lipolytica
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE   1  (bases 1 to 937)
  AUTHORS   Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C.,
            Artiguenave,F., Wincker,P. and Gaillardin,C.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 17. Yarrowia
            lipolytica
  JOURNAL   FEBS Lett. 487 (1), 95-100 (2000)
   PUBMED   11152892
REFERENCE   2  (bases 1 to 937)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
            yeast species for molecular evolution studies(1)
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   PUBMED   11152876
REFERENCE   3  (bases 1 to 937)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This STS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .937
                     /organism="Yarrowia lipolytica"
                     /strain="CLIB 89"
                     /db_xref="taxon:4952"
                     /clone="AW0AA023F10"
                     /clone_lib="AW0AA"
                     /note="end : T3"
     misc_feature    <384. .>701
                     /note="similar to Saccharomyces cerevisiae ORF YNL245c [
                     hypothetical protein ]"
                     /evidence=not_experimental
BASE COUNT      269 a    224 c    251 g    192 t      1 others
ORIGIN


  Query Match             7.1%;  Score 40.6;  DB 53;  Length 937;
  Best Local Similarity   53.5%;  Pred. No. 0.83;
  Matches   85;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     378 CTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCA 437
           ||| || || | |   ||||| ||||||||||| |  |  | |  || |   || | ||
Db     570 CTCGCTTCTTGCCGGTAGCCATCTCCTGCAGCTTTTGAAGCTCGTCTCGTTCACGCTGCT 511

Qy     438 GATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCT 497
           | ||     |||     |   || |   | || ||   |||| |   || ||    ||||
Db     510 GCTTGATCTCCTTCTGGCGCTCCTGGATTCGCCGCCGTTCCTCCAGAGCCTCCTGCTCCT 451

Qy     498 TCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGT 536
           | ||   | ||| | ||| ||   |||     || || |
Db     450 TGGCGTGCACCACGGCTTGGTTTTTGTGCAGACCAGGAT 412
```

```
RESULT  13
HS355C18
LOCUS       HS355C18   101270 bp    DNA             PRI       12-DEC-1999
DEFINITION  Human DNA sequence from clone RP3-355C18 on chromosome 22q13.3
            Contains the KIAA0027 gene, ESTs, STSs, GSSs and seven putative CpG
            islands, complete sequence.
ACCESSION   AL022327
VERSION     AL022327.17  GI:5304851
KEYWORDS    HTG; CpG island; KIAA0027.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 101270)
  AUTHORS   Cobley,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 30, 1999 this sequence version replaced gi:5262834.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP3-355C18 is
            from the library RPCI-3 constructed at the Roswell Park Cancer
            Institute by the group of Pieter de Jong. For further details see
            http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP3-355C18 The true
            left end of clone RP5-898I4 is at 5390 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .101270
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q13.3"
                     /clone="RP3-355C18"
                     /clone_lib="RPCI-3"
     misc_feature    1. .1038
                     /note="match: STS: Em:AL022483; match: STS: Em:HS355C18T"
     repeat_region   1. .242
                     /note="AluJo repeat: matches 1. .237 of consensus"
     repeat_region   247. .322
                     /note="L1 repeat: matches 3067. .3140 of consensus"
     repeat_region   358. .484
                     /note="L1M4c repeat: matches 816. .940 of consensus"
     repeat_region   485. .792
                     /note="AluSx repeat: matches 2. .305 of consensus"
     repeat_region   795. .1095
                     /note="AluSx repeat: matches 1. .299 of consensus"
     repeat_region   1096. .1163
                     /note="L1M4c repeat: matches 940. .1017 of consensus"
     repeat_region   1165. .1551
                     /note="MSTB repeat: matches 3. .412 of consensus"
     repeat_region   1564. .1750
                     /note="L1M4c repeat: matches 1012. .1190 of consensus"
     repeat_region   1757. .2038
                     /note="AluSx repeat: matches 1. .282 of consensus"
     repeat_region   2054. .3342
                     /note="SVA repeat: matches 1. .1386 of consensus"
     repeat_region   3343. .3412
                     /note="35 copies 2 mer tg 82 conserved"
     repeat_region   3418. .3577
                     /note="L1M4c repeat: matches 1190. .1350 of consensus"
     repeat_region   3578. .3861
                     /note="AluJo repeat: matches 1. .300 of consensus"
     repeat_region   3862. .4334
                     /note="L1M4c repeat: matches 1350. .1817 of consensus"
     repeat_region   4341. .4489
                     /note="L1M4c repeat: matches 2052. .1812 of consensus"
     repeat_region   4492. .4858
                     /note="MER4A repeat: matches 237. .660 of consensus"
     repeat_region   5142. .5300
                     /note="AluJb repeat: matches 137. .295 of consensus"
     repeat_region   5315. .5623
                     /note="AluSx repeat: matches 3. .312 of consensus"
     repeat_region   5626. .5786
                     /note="L2 repeat: matches 2492. .2665 of consensus"
     repeat_region   11158. .11247
                     /note="45 copies 2 mer tg 68 conserved"
     repeat_region   11251. .11320
                     /note="35 copies 2 mer tg 72 conserved"
     repeat_region   11369. .11690
                     /note="161 copies 2 mer tg 73 conserved"
     repeat_region   11372. .11679
                     /note="7 copies 44 mer 72 conserved"
     misc_feature    complement(11521)
                     /note="Tandem repeat. Assembly not confirmed by digest"
     misc_feature    11973. .13056
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   12226. .12498
                     /note="13 copies 21 mer 63 conserved"
     repeat_region   14952. .15097
                     /note="MER51B repeat: matches 443. .601 of consensus"
     repeat_region   15195. .15503
                     /note="AluSx repeat: matches 1. .311 of consensus"
     repeat_region   15540. .15948
                     /note="L1MEc repeat: matches 292. .736 of consensus"
     repeat_region   15952. .16252
                     /note="AluSx repeat: matches 1. .306 of consensus"
     repeat_region   16253. .16297
                     /note="Alu repeat: matches 1. .48 of consensus"
     repeat_region   16322. .16780
                     /note="L1M4 repeat: matches 3883. .4321 of consensus"
     repeat_region   16781. .17050
                     /note="AluSx repeat: matches 38. .307 of consensus"
     repeat_region   17051. .17106
                     /note="L1M4 repeat: matches 4321. .4378 of consensus"
     repeat_region   17309. .17576
                     /note="AluSx repeat: matches 22. .298 of consensus"
     repeat_region   17656. .17701
                     /note="23 copies 2 mer aa 80 conserved"
     repeat_region   18229. .18553
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   19381. .19680
                     /note="AluJo repeat: matches 1. .299 of consensus"
     repeat_region   19700. .19923
                     /note="L1 repeat: matches 3883. .4106 of consensus"
     repeat_region   20070. .20379
                     /note="AluSx repeat: matches 9. .310 of consensus"
     repeat_region   20416. .20470
                     /note="U6 repeat: matches 48. .102 of consensus"
     repeat_region   20666. .20821
                     /note="AluSq/x repeat: matches 1. .137 of consensus"
     repeat_region   21593. .21916
                     /note="L2 repeat: matches 2435. .2724 of consensus"
     misc_feature    22848. .23801
                     /note="CpG island"
```

```
                /evidence=not_experimental
     repeat_region   23480. .23537
                /note="29 copies 2 mer gg 72 conserved"
     repeat_region   23650. .23804
                /note="5 copies 31 mer 66 conserved"
     repeat_region   23654. .24010
                /note="17 copies 21 mer 58 conserved"
     repeat_region   26256. .26317
                /note="31 copies 2 mer gt 98 conserved"
     misc_feature    27337. .28417
                /note="CpG island"
                /evidence=not_experimental
     repeat_region   28029. .28422
                /note="197 copies 2 mer cc 56 conserved"
     repeat_region   28038. .28436
                /note="19 copies 21 mer 56 conserved"
     repeat_region   28606. .28661
                /note="28 copies 2 mer cc 71 conserved"
     repeat_region   30789. .30970
                /note="FRAM repeat: matches -8. .174 of consensus"
     repeat_region   30989. .31124
                /note="L1M4c repeat: matches 1258. .1422 of consensus"
     repeat_region   31187. .31308
                /note="L1M4c repeat: matches 1525. .1649 of consensus"
     misc_feature    join(31289. .31500,43139. .43315)
                /note="match: STS: Em:AI026623"
     repeat_region   31510. .32051
                /note="L1M4 repeat: matches 1881. .2118 of consensus"
     repeat_region   32054. .32534
                /note="MLT1F repeat: matches 1. .536 of consensus"
     repeat_region   32541. .32908
                /note="L1M4 repeat: matches 2129. .2499 of consensus"
     repeat_region   33133. .33264
                /note="3 copies 44 mer 74 conserved"
     repeat_region   33279. .33410
                /note="3 copies 44 mer 85 conserved"
     misc_feature    complement(33332)
                /note="Tandem repeat Forced join Each element 22 bases,
                typical sequence: GACAACGTGCACTCCTCCATCT. Restriction
                digest data (ECORI, HINDIII) suggest around 4 copies of
                the repeat are missing from this assembly."
     repeat_region   34730. .34817
                /note="2 copies 44 mer 86 conserved"
     repeat_region   34844. .34931
                /note="2 copies 44 mer 95 conserved"
     repeat_region   35030. .35117
                /note="2 copies 44 mer 95 conserved"
     repeat_region   35455. .35648
                /note="L1M4 repeat: matches 4149. .4351 of consensus"
     repeat_region   35733. .36111
                /note="L1M4 repeat: matches 4263. .4643 of consensus"
     repeat_region   36112. .36420
                /note="AluSc repeat: matches 1. .307 of consensus"
     repeat_region   36421. .36447
                /note="L1M4 repeat: matches 4643. .4669 of consensus"
     repeat_region   36448. .36756

  Query Match             7.0%;  Score 40;  DB 92;  Length 101270;
  Best Local Similarity   61.5%;  Pred. No. 1;
  Matches   64;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy      36 GCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCA 95
           |||| || || |   ||  |  | |     ||| ||  |    ||||||||| || ||||
Db   61606 GCTGGCCAGGCCCCATCCTGCAACACACCAGCCGGGCTGTCCGAGCCCTGCCGGGCCTCA 61665

Qy      96 GTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTG 139
           |||||  |||||| ||||   || | |||  |  |||  | |||
Db   61666 GTGCCCTCAGCTTCCACATGCGCACAGCAGGCCAGCCCAGTGTG 61709


RESULT  14
AC022220
LOCUS       AC022220   233231 bp    DNA             HTG       17-AUG-2000
DEFINITION  Homo sapiens chromosome X clone RP11-815F17, WORKING DRAFT
            SEQUENCE, 10 unordered pieces.
ACCESSION   AC022220
VERSION     AC022220.7  GI:9838350
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 233231)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 233231)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 17, 2000 this sequence version replaced gi:8980041.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0815F17
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 66%
            Sequencing vector: plasmid; 34%
            Chemistry: Dye-primer ET; 66% of reads
            Chemistry: Dye-terminator Big Dye; 34% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 228369 bases at least Q40
            Consensus quality: 229897 bases at least Q30
            Consensus quality: 230504 bases at least Q20
            Insert size: 214000; agarose-fp
            Insert size: 233559; sum-of-contigs
            Quality coverage: 7.60 in Q20 bases; agarose-fp
            Quality coverage: 6.98 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1224: contig of 1224 bp in length
            *     1225     1324: gap of unknown length
            *     1325    11218: contig of 9894 bp in length
            *    11219    11318: gap of unknown length
            *    11319    20516: contig of 9198 bp in length
            *    20517    20616: gap of unknown length
            *    20617    34259: contig of 13643 bp in length
            *    34260    34359: gap of unknown length
            *    34360    47343: contig of 12984 bp in length
            *    47344    47443: gap of unknown length
            *    47444    63537: contig of 16094 bp in length
            *    63538    63637: gap of unknown length
            *    63638    91959: contig of 28322 bp in length
            *    91960    92059: gap of unknown length
            *    92060   127450: contig of 35391 bp in length
            *   127451   127550: gap of unknown length
            *   127551   182595: contig of 55045 bp in length
            *   182596   182695: gap of unknown length
            *   182696   233231: contig of 50536 bp in length.
FEATURES             Location/Qualifiers
     source          1. .233231
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
```

```
                     /clone="RP11-815F17"
     misc_feature    1. .1224
                     /note="assembly_name:Contig29"
     misc_feature    1325. .11218
                     /note="assembly_name:Contig30"
     misc_feature    11319. .20516
                     /note="assembly_name:Contig31"
     misc_feature    20617. .34259
                     /note="assembly_name:Contig32"
     misc_feature    34360. .47343
                     /note="assembly_name:Contig33"
     misc_feature    47444. .63537
                     /note="assembly_name:Contig34
                     clone_end:SP6
                     vector_side:right"
     misc_feature    63638. .91959
                     /note="assembly_name:Contig35"
     misc_feature    92060. .127450
                     /note="assembly_name:Contig36
                     clone_end:T7
                     vector_side:right"
     misc_feature    127551. .182595
                     /note="assembly_name:Contig37"
     misc_feature    182696. .233231
                     /note="assembly_name:Contig38"
BASE COUNT    64468 a  48611 c  49543 g  69686 t    923 others
ORIGIN


  Query Match             6.9%;  Score 39.4;  DB 67;  Length 233231;
  Best Local Similarity   67.9%;  Pred. No. 1.5;
  Matches   55;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy    329 CAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGG 388
          ||| || ||  |||| |||| |||  ||| |||   | | | |||| |||   || |||
Db 133397 CAGTGGCGCGATCTCGGCTCACTGCAAGCTCTGCCTCCCGGGTTCACGCCATTCTCCTGT 133456

Qy    389 GCCTCAGCCAACTCCTGCAGC 409
          |||||||||  | || | |||
Db 133457 GCCTCAGCCTCCCCCAGTAGC 133477


RESULT  15
AC011100/c
LOCUS       AC011100    55815 bp    DNA             HTG       26-JAN-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-18F6 map 15, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC011100
VERSION     AC011100.3  GI:12545324
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55815)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-18F6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 55815)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 26, 2001 this sequence version replaced gi:8901225.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1669
                Center clone name: 18_F_6
            ---------------.
            * NOTE: This record contains 68 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      741: contig of 741 bp in length
            *      742 841: gap of      100 bp
            *      842     1566: contig of 725 bp in length
            *     1567 1666: gap of      100 bp
            *     1667     2382: contig of 716 bp in length
            *     2383 2482: gap of      100 bp
            *     2483     3203: contig of 721 bp in length
            *     3204 3303: gap of      100 bp
            *     3304     4022: contig of 719 bp in length
            *     4023 4122: gap of      100 bp
            *     4123     4850: contig of 728 bp in length
            *     4851 4950: gap of      100 bp
            *     4951     5684: contig of 734 bp in length
            *     5685 5784: gap of      100 bp
            *     5785     6509: contig of 725 bp in length
            *     6510 6609: gap of      100 bp
            *     6610     7352: contig of 743 bp in length
            *     7353 7452: gap of      100 bp
            *     7453     8190: contig of 738 bp in length
            *     8191 8290: gap of      100 bp
            *     8291     9012: contig of 722 bp in length
            *     9013 9112: gap of      100 bp
            *     9113     9819: contig of 707 bp in length
            *     9820 9919: gap of      100 bp
            *     9920    10624: contig of 705 bp in length
            *    10625 10724: gap of      100 bp
            *    10725    11468: contig of 744 bp in length
            *    11469 11568: gap of      100 bp
            *    11569    12301: contig of 733 bp in length
            *    12302 12401: gap of      100 bp
            *    12402    13133: contig of 732 bp in length
            *    13134 13233: gap of      100 bp
            *    13234    13954: contig of 721 bp in length
            *    13955 14054: gap of      100 bp
            *    14055    14668: contig of 614 bp in length
            *    14669 14768: gap of      100 bp
            *    14769    15499: contig of 731 bp in length
            *    15500 15599: gap of      100 bp
            *    15600    16341: contig of 742 bp in length
            *    16342 16441: gap of      100 bp
            *    16442    17169: contig of 728 bp in length
            *    17170 17269: gap of      100 bp
            *    17270    17991: contig of 722 bp in length
            *    17992 18091: gap of      100 bp
            *    18092    18802: contig of 711 bp in length
            *    18803 18902: gap of      100 bp
```

```
*    18903    19615: contig of 713 bp in length
*    19616 19715: gap of      100 bp
*    19716    20445: contig of 730 bp in length
*    20446 20545: gap of      100 bp
*    20546    21269: contig of 724 bp in length
*    21270 21369: gap of      100 bp
*    21370    22091: contig of 722 bp in length
*    22092 22191: gap of      100 bp
*    22192    22934: contig of 743 bp in length
*    22935 23034: gap of      100 bp
*    23035    23792: contig of 758 bp in length
*    23793 23892: gap of      100 bp
*    23893    24630: contig of 738 bp in length
*    24631 24730: gap of      100 bp
*    24731    25469: contig of 739 bp in length
*    25470 25569: gap of      100 bp
*    25570    26289: contig of 720 bp in length
*    26290 26389: gap of      100 bp
*    26390    27140: contig of 751 bp in length
*    27141 27240: gap of      100 bp
*    27241    27946: contig of 706 bp in length
*    27947 28046: gap of      100 bp
*    28047    28769: contig of 723 bp in length
*    28770 28869: gap of      100 bp
*    28870    29551: contig of 682 bp in length
*    29552 29651: gap of      100 bp
*    29652    30396: contig of 745 bp in length
*    30397 30496: gap of      100 bp
*    30497    31219: contig of 723 bp in length
*    31220 31319: gap of      100 bp
*    31320    32021: contig of 702 bp in length
*    32022 32121: gap of      100 bp
*    32122    32842: contig of 721 bp in length
*    32843 32942: gap of      100 bp
*    32943    33656: contig of 714 bp in length
*    33657 33756: gap of      100 bp
*    33757    34468: contig of 712 bp in length
*    34469 34568: gap of      100 bp
*    34569    35295: contig of 727 bp in length
*    35296 35395: gap of      100 bp
*    35396    36129: contig of 734 bp in length
*    36130 36229: gap of      100 bp
*    36230    36949: contig of 720 bp in length
*    36950 37049: gap of      100 bp
*    37050    37790: contig of 741 bp in length
*    37791 37890: gap of      100 bp
*    37891    38608: contig of 718 bp in length
*    38609 38708: gap of      100 bp
*    38709    39414: contig of 706 bp in length
*    39415 39514: gap of      100 bp
*    39515    40251: contig of 737 bp in length
*    40252 40351: gap of      100 bp
*    40352    41047: contig of 696 bp in length
*    41048 41147: gap of      100 bp
*    41148    41860: contig of 713 bp in length
*    41861 41960: gap of      100 bp
*    41961    42695: contig of 735 bp in length
*    42696 42795: gap of      100 bp
*    42796    43493: contig of 698 bp in length
*    43494 43593: gap of      100 bp
*    43594    44347: contig of 754 bp in length
*    44348 44447: gap of      100 bp
*    44448    45176: contig of 729 bp in length
*    45177 45276: gap of      100 bp
*    45277    45986: contig of 710 bp in length
*    45987 46086: gap of      100 bp
*    46087    46806: contig of 720 bp in length
*    46807 46906: gap of      100 bp
*    46907    47627: contig of 721 bp in length
*    47628 47727: gap of      100 bp
*    47728    48445: contig of 718 bp in length
*    48446 48545: gap of      100 bp
*    48546    49267: contig of 722 bp in length
*    49268 49367: gap of      100 bp
*    49368    50091: contig of 724 bp in length
*    50092 50191: gap of      100 bp
*    50192    50922: contig of 731 bp in length
*    50923 51022: gap of      100 bp
*    51023    51712: contig of 690 bp in length
*    51713 51812: gap of      100 bp
*    51813    52519: contig of 707 bp in length
*    52520 52619: gap of      100 bp
*    52620    53336: contig of 717 bp in length
*    53337 53436: gap of      100 bp
*    53437    54145: contig of 709 bp in length
*    54146 54245: gap of      100 bp
*    54246    54976: contig of 731 bp in length
*    54977 55076: gap of      100 bp
*    55077    55815: contig of 739 bp in length.
FEATURES             Location/Qualifiers
     source          1. .55815
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-18F6"
                     /clone_lib="RPCI-11 Human Male BAC"
BASE COUNT    13836 a   9811 c  10383 g  14691 t   7094 others

  Query Match              6.8%;  Score 39;  DB 62;  Length 55815;
  Best Local Similarity   54.5%;  Pred. No. 2;
  Matches   78;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy     348 CCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCA 407
           ||| | | |  |   |   ||||| | ||   |  | ||| | ||| |||| || ||| |
Db    8731 CCCAGCTGGGGCAACGCACCAGCTCCTTGAAGCTGTTCTGTGGCTCGGCCAGCTGCTGAA 8672

Qy     408 GCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATG 467
           ||||       |  | |||| || |||  | | | |          || |  || ||
Db    8671 GCCTCTCCTCCTGCTCTGGAAGCCTCTCCTGCTGCTCCTGAAGCCTCTCCTCCTGCTCCC 8612

Qy     468 GCAGCGTCTCCTTCTCCGCTTCA 490
           | ||| ||||||| | | | |||
Db    8611 GAAGCCTCTCCTTTTGCCCCTCA 8589


Search completed: June 21, 2001, 02:07:38
Job time: 8327 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2001, 00:52:31 ; Search time 153.2 Seconds
(without alignments)
2172.032 Million cell updates/sec

Title: US-09-558-474-1
Perfect score: 570
Sequence: 1 ATGCTGGGGAGCAGAGCTGT.....CAGCAACCCTGAGTCCCTAA 570

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0401:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	570	100.0	570	20	Z08865	Human interleukin
2	570	100.0	570	20	X17786	Human interleukin-
3	570	100.0	570	21	A52577	Human interleukin-
4	566.8	99.4	1026	21	Z37262	SGRF coding sequen
5	566.8	99.4	1055	22	C84306	Human EXCS encodin
6	566.8	99.4	1067	21	A16686	Human secreted pro
7	391	68.6	1203	20	Z08866	Mouse interleukin
8	391	68.6	1203	20	X17787	Mouse interleukin-
9	391	68.6	1203	21	A52578	Murine interleukin
10	242.8	42.6	412	21	A42676	Human secreted exp
11	192.8	33.8	2398	21	Z37263	SGRF coding sequen
c 12	38	6.7	121162	21	C66548	Human kinesin-like
13	35	6.1	567	21	Z55567	Feline interferon-
c 14	35	6.1	567	21	Z55568	Feline interferon-
15	35	6.1	567	21	Z55569	Feline interferon-
c 16	35	6.1	567	21	Z55570	Feline interferon-
17	35	6.1	582	10	N90186	Feline interferon
18	35	6.1	927	18	T51317	HSV glycoprotein D
19	35	6.1	1185	20	Z10959	HSV-1 glycoprotein
20	35	6.1	1550	18	T51318	HSV glycoprotein D
21	35	6.1	1608	5	N40070	Sequence coding fo
22	35	6.1	1608	7	N60545	Sequence of a port
23	34.8	6.1	1594	6	N50489	Sequence of herpes
24	34.8	6.1	1602	10	N90524	Herpes Simplex vir
c 25	34.6	6.1	985	21	A64363	Mesothelin related
c 26	34.6	6.1	1198	21	A64362	Mesothelin related
c 27	34.6	6.1	1207	21	A64366	Soluble mesothelin
c 28	34.6	6.1	2129	15	Q63972	Megakaryoctye pote
c 29	34.6	6.1	2138	18	T91079	Human CAK1 antigen
c 30	34.2	6.0	4308	20	Z34327	Human EST DNA34415
c 31	34.2	6.0	4308	21	C78593	Human EST DNA34415
c 32	33.8	5.9	1203	21	A59324	Coding region of a
c 33	33.8	5.9	1632	19	V68059	Neurodegenerative
c 34	33.8	5.9	2516	21	A59323	DNA encoding a hum
c 35	33.8	5.9	2711	19	V68056	Neurodegenerative
36	33.4	5.9	694	21	A42007	Human secreted exp
37	33.4	5.9	1218	12	Q11073	HSV-1 antigen/heat
38	33.4	5.9	1308	12	Q10209	Fusion gene compri
39	33.4	5.9	1308	12	Q11074	HSV-1 antigen/heat
40	33.4	5.9	1459	11	Q05677	Sequence encoding
41	33.4	5.9	2925	12	Q14455	HSV surface antige
c 42	33	5.8	1051	19	V27386	Streptococcus pneu
43	33	5.8	1319	18	X30940	Streptococcus pneu
c 44	33	5.8	5052	21	C50830	Arabidopsis thalia
45	33	5.8	13206	19	V52166	Streptococcus pneu

ALIGNMENTS

RESULT 1
Z08865
ID Z08865 standard; cDNA; 570 BP.
XX
AC Z08865;
XX
DT 04-NOV-1999 (first entry)
XX
DE Human interleukin B30 encoding cDNA.
XX
KW DNAX soluble receptor subunit 1; DNAX cytokine receptor subunit 1;
KW interleukin B30; DSRS1; DCRS1; IL-B30; cytokine receptor; diagnosis;
KW inflammatory disorder; inflammatory response; innate immunity;
KW morphogenic development; immunological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..570
FT /*tag= a
FT /product= "IL-B30"
FT sig_peptide 1..63
FT /*tag= b
FT mat_peptide 64..567
FT /*tag= c
XX
PN WO9940195-A1.
XX
PD 12-AUG-1999.
XX
PF 05-FEB-1999; 99WO-US02600.
XX
PR 13-MAY-1998; 98US-0078194.
PR 06-FEB-1998; 98US-0073941.

```
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Kastelein RA,  Mattson JD,  McClanahan TK;
XX
DR   WPI; 1999-527306/44.
DR   P-PSDB; Y29783.
XX
PT   New receptor subunits useful in the treatment inflammatory disorders
XX
PS   Claim 16; Page 26-27; 133pp; English.
XX
CC   The present invention describes a composition (I) comprising DNAX
CC   cytokine receptor subunit I (DCRS1) protein and DNAX soluble receptor
CC   subunit I (DSRS1) protein, which together encode a new mammalian
CC   cytokine-related receptor (R), or DCRS1 and interleukin B30 (IL-B30)
CC   proteins, or DSRS1 and IL-B30 proteins. (I) comprising DSRS1 and DCRS1
CC   is useful for screening for ligands (i.e. agonists/antagonists) from
CC   a library of compounds, which are useful for modulating the physiology
CC   or development of a cell or tissue culture e.g. inflammatory responses,
CC   innate immunity and/or morphogenic development. (R), antibodies and
CC   ligands are useful for treatment of conditions, especially immunological
CC   disorders, associated with conditions exhibiting abnormal expression of
CC   (R). (R) is useful as a phosphate labeling enzyme to label substrates,
CC   and the subunits DSRS1 and DCRS1 are useful as immunogens for generating
CC   antibodies, or as antigens for binding antibodies. Nucleic acids
CC   encoding (R) are useful for identifying related DNAs and mRNAs, and
CC   variants from other individuals or species. The present sequence
CC   encodes the specifically claimed human IL-B30, for use in the
CC   composition of the present invention.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;

  Query Match             100.0%;  Score 570;  DB 20;  Length 570;
  Best Local Similarity   100.0%;  Pred. No. 2e-157;
  Matches  570;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gctgtgcctggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 180

Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgaccccaa 240

Qy    241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttat 300

Qy    301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 360

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 420

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 480

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 540
```

```
Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    541 gcccatggagcagcaaccctgagtccctaa 570
```

Name invention

```
RESULT 2
X17786
ID   X17786 standard; cDNA; 570 BP.
XX
AC   X17786;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human interleukin-B30 (IL-B30) polypeptide encoding cDNA.
XX
KW   Cytokine; interleukin-B30; IL-B30; forensic science; cell proliferation;
KW   inflammatory condition; drug screening; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..570
FT                   /*tag= a
FT                   /product= "IL-B30"
FT   sig_peptide     1..63
FT                   /*tag= b
FT   mat_peptide     64..567
FT                   /*tag= c
XX
PN   WO9905280-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;   98WO-US15423.
XX
PR   25-JUL-1997;   97US-0900905.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 1999-142935/12.
DR   P-PSDB; W95002.
XX
PT   Newly isolated or recombinant polynucleotide encoding mammalian
PT   cytokine interleukin-B30 (IL-B30), including fragments - useful for
PT   regulating activation, development, differentiation and function of
PT   various cell types, and for diagnosing and treating conditions
PT   associated with IL-B30
XX
PS   Claim 3; Page 8-9; 83pp; English.
XX
CC   This cDNA encodes a human cytokine interleukin-B30 (IL-B30) polypeptide.
CC   Host cells containing a vector comprising the IL-B30 nucleic acid are
CC   used for the recombinant production of the protein. The polynucleotides
CC   are useful for diagnosis of IL-B30 mediated conditions, and forensic
CC   science (e.g. to distinguish rodent from human, or as a marker to
CC   distinguish between different cells exhibiting differential expression or
CC   modification patterns). The IL-B30 (including fragments), together with
CC   antibodies that bind to IL-B30 are useful for teaching purposes. They
CC   are also used for treating conditions associated with abnormal physiology
CC   or development, including inflammatory conditions. The polypeptide
CC   cytokine should mediate cytokine synthesis and proliferation in cells.
CC   IL-B30 is useful for drug screening to identify compounds having binding
CC   affinity to IL-B30.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;

  Query Match             100.0%;  Score 570;  DB 20;  Length 570;
  Best Local Similarity   100.0%;  Pred. No. 2e-157;
  Matches  570;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 180

Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgacccccaa 240

Qy    241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttat 300

Qy    301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 360

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 420

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 480

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 540

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          |||||||||||||||||||||||||||||
Db    541 gcccatggagcagcaaccctgagtccctaa 570
```

```
RESULT   3
A52577
ID   A52577 standard; cDNA; 570 BP.
XX
AC   A52577;
XX
DT   27-SEP-2000  (first entry)
XX
DE   Human interleukin-B30 (IL-B30) cDNA.
XX
KW   Interleukin-B30; IL-B30; human; cytokine; cellular signalling;
KW   immune response; haematopoietic cell; IL-6 homologue; GCSF homologue;
KW   granulocyte colony stimulating factor; immune disorder;
KW   inflammatory disease; autoimmune disease; antigen; antibody; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..570
FT                   /*tag=  a
FT                   /product= "Human IL-B30 precursor"
FT   sig_peptide     1..63
FT                   /*tag=  b
FT   mat_peptide     64..567
FT                   /*tag=  c
FT                   /product= "Mature human IL-B30"
XX
PN   US6060284-A.
XX
PD   09-MAY-2000.
XX
PF   24-JUL-1998;   98US-0122443.
XX
PR   25-JUL-1997;   97US-0053765.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 2000-364420/31.
DR   P-PSDB; B01981.
XX
PT   Novel recombinant DNA encoding cytokines especially interleukin-B30
PT   useful as probes or primers for diagnosing immune disorders including
PT   autoimmune or chronic inflammatory conditions  -
XX
PS   Claim 3; Column 5-6; 32pp; English.
XX
CC   This sequence represents cDNA encoding human interleukin-B30 (IL-B30).
CC   IL-B30 is a novel cytokine, exhibiting significant homology to IL-6 and
CC   GCSF (granulocyte colony stimulating factor). Cytokines play a critical
CC   role in signalling between immune or other cells during an immune
CC   response. The precise role if IL-B30 is not yet known - it is likely to
CC   have either a stimulatory or an inhibitory effect on haematopoietic cells
CC   such as T-cells, B-cells, natural killer (NK) cells and macrophages.
CC   Alternatively, it may affect vascular physiology or development, or have
CC   neuronal effects. IL-B30, its fragments, IL-B30 nucleotides, agonists and
CC   antagonists are useful in the diagnosis and treatment of disorders
CC   associated with abnormal expression or activity of IL-B30 e.g.,
CC   inflammatory or autoimmune disorders. Nucleotides encoding IL-B30 are
CC   useful for recombinant expression of IL-B30 in a host cell, and as a
CC   source of probes and primers. The IL-B30 probes and primers can be used
CC   to detect levels of IL-B30 expression in samples from patients suspected
CC   of having an inflammatory or autoimmune disorder. IL-B30 nucleotides may
CC   also be used to identify homologous genes in other species. IL-B30
CC   protein or its fragments are useful as antigens for raising antibodies to
CC   various linear and conformational epitopes. Such antibodies may be used
CC   to detect levels of IL-B30 protein in a sample.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;
```

```
  Query Match             100.0%;  Score 570;  DB 21;  Length 570;
  Best Local Similarity   100.0%;  Pred. No. 2e-157;
  Matches  570;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 180

Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgacccccaa 240

Qy    241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttat 300

Qy    301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 360

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 420
```

```
Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 480

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 540

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    541 gcccatggagcagcaaccctgagtccctaa 570


RESULT   4
Z37262
ID   Z37262 standard; DNA; 1026 BP.
XX
AC   Z37262;
XX
DT   01-FEB-2000  (first entry)
XX
DE   SGRF coding sequence.
XX
KW   SGRF; human; Interleukin-6 G-CSF related factor; cell proliferation;
KW   immune system; haematopoietic system; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9954357-A1.
XX
PD   28-OCT-1999.
XX
PF   14-APR-1999;   99WO-JP01997.
XX
PR   14-APR-1998;   98JP-0121805.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Hirata Y;
XX
DR   WPI; 2000-013230/01.
DR   P-PSDB; Y54606.
XX
PT   Novel cytokine-like protein, with activity of supporting proliferation
PT   of myeloid cells, useful in treating abnormality of cell proliferation
PT   in immune and haematopoiesis systems -
XX
PS   Claim 2; Fig 1; 69pp; Japanese.
XX
CC   This sequence encodes the Interleukin-6 G-CSF related factor (SGRF)
CC   protein of the invention. The protein is a member of the IL-6/G-CSF/MSF
CC   family. The protein can be used in drugs for treating diseases due to
CC   abnormality of cell proliferation in the immune system and haematopoietic
CC   system.
XX
SQ   Sequence 1026 BP; 268 A; 249 C; 272 G; 237 T; 0 other;


  Query Match             99.4%;  Score 566.8;  DB 21;  Length 1026;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-156;
  Matches  568;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 203

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 263

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 323

Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgaccccaa 383

Qy    241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    384 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttat 443

Qy    301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    444 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 503

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504 gtgggccagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 563

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    564 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 623

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    624 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 683

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    684 gcccatggagcagcaaccctgagtccctaa 713


RESULT   5
C84306
ID   C84306 standard; cDNA; 1055 BP.
XX
AC   C84306;
XX
DT   19-MAR-2001  (first entry)
XX
DE   Human EXCS encoding cDNA (clone ID 2933038CB1).
XX
KW   Extracellular signaling molecule; EXCS; anti-inflammatory; human;
KW   immunosuppressive; cytostatic; neuroprotective; gastrointestinal;
KW   virucide; antibacterial; anti-HIV; human immunodeficiency virus;
KW   antiinfertility; cerebroprotective; nootropic; antiulcer; antifungal;
KW   anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological;
KW   keratolytic; protozoacide; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200070049-A2.
XX
PD   23-NOV-2000.
XX
PF   19-MAY-2000; 2000WO-US13975.
XX
PR   19-MAY-1999;   99US-0134949.
PR   15-JUL-1999;   99US-0144270.
PR   30-JUL-1999;   99US-0146700.
PR   04-OCT-1999;   99US-0157508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Lal P,  Burford N,  Bandman O,  Baughn MR;
PI   Azimzai Y,  Lu DAM,  Patterson C;
XX
DR   WPI; 2001-025021/03.
DR   P-PSDB; B48070.
XX
PT   New human extracellular signaling nucleic acids and polypeptides useful
PT   for diagnosing, treating and preventing infections and
PT   gastrointestinal, neurological, reproductive, and
PT   autoimmune/inflammatory disorders -
XX
```

```
PS   Claim 4; Page 106; 114pp; English.
XX
CC   The invention provides human extracellular signaling molecules (EXCS)
CC   and polynucleotides which identify and encode EXCS. EXCS can be expressed
CC   by standard recombinant methodology. The amino acid and nucleic acid
CC   sequences of EXCS are useful for diagnosing, treating and preventing
CC   infections and gastrointestinal (peptic ulcer, dysphagia, pancreatitis),
CC   neurological (e.g. epilepsy, ischemic cerebrovascular disease, stroke),
CC   reproductive (infertility, ovulatory defects, endometriosis), autoimmune
CC   /inflammatory (actinic keratosis, acquired immunodeficiency syndrome
CC   (AIDS), Addison's disease), and cell proliferative disorders including
CC   cancers (of the breast, adrenal gland, bone). They may also be used to
CC   treat fatal familial insomnia, nutritional and metabolic diseases of the
CC   nervous system, myopathies, mental disorders (anxiety, schizophrenia,
CC   mood), as well as infections caused by parasites (malaria, leishmania,
CC   trypanosoma), viral (adenovirus, coronavirus, flavivirus), bacterial
CC   (e.g. pneumococcus, staphylococcus, bacillus), and fungal (aspergillus,
CC   blastomyces, dermatophytes) agents. The nucleic acids, polypeptides,
CC   antagonists, agonists, pharmaceutical compositions, and antibodies may
CC   also be used for treating or preventing disorders associated with
CC   increased or decreased expression or activity of EXCS. EXCS
CC   polynucleotides may also be used to detect and quantify gene expression
CC   in biopsied tissues in which expression of EXCS may be correlated with
CC   the disease, to determine presence or excess expression of EXCS, to
CC   monitor regulation of EXCS levels during therapeutic intervention, to
CC   detect the presence of associated disorders, as targets in microarray, to
CC   generate hybridization probes, and to detect differences in gene
CC   sequences among normal, carrier or affected individuals. Antibodies may
CC   also be used in diagnosing disorders, in monitoring patients being
CC   treated with EXCS agonists, antagonists or inhibitors. Sequences
CC   C84293-C84318 represent nucleic acid molecules encoding the EXCS of the
CC   invention.
XX
SQ   Sequence 1055 BP; 284 A; 254 C; 278 G; 239 T; 0 other;


  Query Match             99.4%;  Score 566.8;  DB 22;  Length 1055;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-156;
  Matches  568;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 230

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 290

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 350

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgacccccaa 410

Qy     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttat 470

Qy     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 530

Qy     361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
           ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     531 gtgggccagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 590

Qy     421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     591 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 650

Qy     481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     651 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 710

Qy     541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           |||||||||||||||||||||||||||||
Db     711 gcccatggagcagcaaccctgagtccctaa 740


RESULT   6
A16686
ID   A16686 standard; cDNA; 1067 BP.
XX
AC   A16686;
XX
DT   16-JUN-2000  (first entry)
XX
DE   Human secreted protein clone rk80_3 nucleotide sequence SEQ ID NO:137.
XX
KW   Human; secreted protein; immunestimulant; immunesuppressant; virucide;
KW   antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
KW   antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;
KW   antithyroid; immune deficiency; severe combined immunedeficiency; SCID;
KW   infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
KW   connective tissue disease; multiple sclerosis; erythematosis;
KW   rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
KW   Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
KW   insulin dependent diabetes mellitus; graft-versus-host-disease;
KW   autoimmune inflammatory eye disease; allergy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009552-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1999;   99WO-US18298.
XX
PR   14-AUG-1998;   98US-0096622.
PR   17-AUG-1998;   98US-0096815.
PR   04-SEP-1998;   98US-0099229.
PR   23-OCT-1998;   98US-0105368.
PR   08-JAN-1999;   99US-0115234.
PR   12-FEB-1999;   99US-0119931.
PR   18-FEB-1999;   99US-0120575.
PR   30-APR-1999;   99US-0132020.
PR   11-AUG-1999;   99US-0096622.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M,  Agostino MJ,  Steininger RJ,  Spaulding V;
PI   Wong GG,  Clark HF,  Fechtel K;
XX
DR   WPI; 2000-205979/18.
DR   P-PSDB; Y94966.
XX
PT   New polynucleotides encoding secreted proteins, which may have e.g.
PT   nutritional, chemokine, immune stimulating or suppressing,
PT   hematopoiesis regulating, tissue growth, activin/inhibin
PT   antiinflammatory or tumor inhibition activity  -
XX
PS   Claim 146; Page 596-597; 641pp; English.
XX
CC   A16618 to A16697 encode the human secreted proteins given in Y94898
CC   to Y94980, isolated from human adult brain, adult thyroid, adult retina,
CC   foetal carcinoma, adult blood, adult neural, foetal kidney, adult
CC   placenta, adult testis, whole embryo, adult cartilage, kidney, foetal
CC   brain, adult thymus, foetal placenta, adult uterus, adult tumour, and
CC   adult bladder, cDNA libraries. The polynucleotides and proteins are
CC   predicted to have biological activities which would make them suitable
CC   for treating, preventing or ameliorating medical conditions in humans
CC   and animals. The polynucleotides can be used as markers for tissues in
CC   which the protein is preferentially expressed, as molecular weight
```

```
CC   markers on Southern gels, and as chromosome markers or tags to identify
CC   chromosomes or to map gene positions. The proteins can be used in the
CC   treatment of immune deficiencies and disorders, such as severe combined
CC   immunedeficiency (SCID), as well as viral, bacterial, fungal and other
CC   infections. These infections include human immunodeficiency virus (HIV),
CC   hepatitis, herpesviruses, mycobacteria, Leismania spp., malaria and
CC   candidiasis. The proteins can be used to treat autoimmune disorders such
CC   as connective tissue disease, multiple sclerosis, systemic lupus
CC   erythematosis, rheumatoid arthritis, autoimmune pulmonary inflammation,
CC   Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
CC   diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
CC   autoimmune inflammatory eye disease. The proteins can also be used to
CC   treat allergic conditions, such asthma. A16698 to A16774 represent
CC   probes for the human secreted proteins from the present invention.
XX
SQ   Sequence 1067 BP; 293 A; 255 C; 281 G; 238 T; 0 other;


  Query Match             99.4%;  Score 566.8;  DB 21;  Length 1067;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-156;
  Matches  568;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     179 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 238

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     239 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 298

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 358

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     359 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgaccccaa 418

Qy     241 GGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 ggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgatttttttat 478

Qy     301 GAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 gagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagccct 538

Qy     361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
           ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539 gtgggccagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggtcac 598

Qy     421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     599 cactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctcctt 658

Qy     481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     659 ctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtcttt 718

Qy     541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||
Db     719 gcccatggagcagcaaccctgagtccctaa 748


RESULT   7
Z08866
ID   Z08866 standard; cDNA; 1203 BP.
XX
AC   Z08866;
XX
DT   04-NOV-1999  (first entry)
XX
DE   Mouse interleukin B30 encoding cDNA.
XX
KW   DNAX soluble receptor subunit 1; DNAX cytokine receptor subunit 1;
KW   interleukin B30; DSRS1; DCRS1; IL-B30; cytokine receptor; diagnosis;
KW   inflammatory disorder; inflammatory response; innate immunity;
KW   morphogenic development; immunological disorder; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..703
FT                   /*tag= a
FT                   /product= "IL-B30"
FT   sig_peptide     113..175
FT                   /*tag= b
FT   mat_peptide     176..700
FT                   /*tag= c
XX
PN   WO9940195-A1.
XX
PD   12-AUG-1999.
XX
PF   05-FEB-1999;   99WO-US02600.
XX
PR   13-MAY-1998;   98US-0078194.
PR   06-FEB-1998;   98US-0073941.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Kastelein RA,  Mattson JD,  McClanahan TK;
XX
DR   WPI; 1999-527306/44.
DR   P-PSDB; Y29784.
XX
PT   New receptor subunits useful in the treatment inflammatory disorders
XX
PS   Claim 16; Page 27-28; 133pp; English.
XX
CC   The present invention describes a composition (I) comprising DNAX
CC   cytokine receptor subunit I (DCRS1) protein and DNAX soluble receptor
CC   subunit I (DSRS1) protein, which together encode a new mammalian
CC   cytokine-related receptor (R), or DCRS1 and interleukin B30 (IL-B30)
CC   proteins, or DSRS1 and IL-B30 proteins. (I) comprising DSRS1 and DCRS1
CC   is useful for screening for ligands (i.e. agonists/antagonists) from
CC   a library of compounds, which are useful for modulating the physiology
CC   or development of a cell or tissue culture e.g. inflammatory responses,
CC   innate immunity and/or morphogenic development. (R), antibodies and
CC   ligands are useful for treatment of conditions, especially immunological
CC   disorders, associated with conditions exhibiting abnormal expression of
CC   (R). (R) is useful as a phosphate labeling enzyme to label substrates,
CC   and the subunits DSRS1 and DCRS1 are useful as immunogens for generating
CC   antibodies, or as antigens for binding antibodies. Nucleic acids
CC   encoding (R) are useful for identifying related DNAs and mRNAs, and
CC   variants from other individuals or species. The present sequence
CC   encodes the specifically claimed mouse IL-B30, for use in the
CC   composition of the present invention.
XX
SQ   Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;


  Query Match             68.6%;  Score 391;  DB 20;  Length 1203;
  Best Local Similarity   81.8%;  Pred. No. 7.5e-105;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           |||||||   |||||||  |||||  ||||  |  ||||  ||||||||    ||||||||
Db     113 atgctggattgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctg 172

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           |||||||||  ||  |  |||||  ||||  ||||  ||||||||||||||||  ||  |||  |||
Db     173 gctgtgcctaggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctc 232

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
           ||||  ||  |||||||  ||||||  |||  ||  ||||||  |||  ||   ||||||||||||
Db     233 tgcatgctagcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaa 292
```

```
Qy    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
          |  ||||||||||||| |||| |||| |||| |||||||||| |||||| ||||||||
Db    293 gaggatgaagagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgaccca 352

Qy    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
          |||||||||| |||||||| ||||||||||||||||||||||| ||| ||||||  |||| 
Db    353 caaggactcaaggacaacagccagttctgcttgcaaaggatccgccaaggtctggctttt 412

Qy    298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||| 
Db    413 tataagcacctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagc 472

Qy    358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417
          ||  ||| ||| ||||| |||||||||| || |||||||||||||| ||||| |||| | 
Db    473 cccatggagcaacttcacacctccctactaggactcagccaactcctccagccagaggat 532

Qy    418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477
          |||| | ||||||| || ||||| || ||||| ||| | || |||| ||||||||| | | 
Db    533 caccccgggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgcccc 592

Qy    478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537
          |||||||| | ||| |||||||| |||||||||||||| |||| |||| |||||||||
Db    593 cttctccgttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtc 652

Qy    538 TTTGCCCATGGAGCAGCAACCCTGAGT 564
          |||||||| |||||||||| |||| |
Db    653 tttgcccacggagcagcaactctgact 679
```

RESULT 8
X17787
ID X17787 standard; cDNA; 1203 BP.
XX
AC X17787;
XX
DT 21-MAY-1999 (first entry)
XX
DE Mouse interleukin-B30 (IL-B30) polypeptide encoding cDNA.
XX
KW Cytokine; interleukin-B30; IL-B30; forensic science; cell proliferation;
KW inflammatory condition; drug screening; mouse; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 113..703
FT /*tag= a
FT /product= "IL-B30"
FT sig_peptide 113..175
FT /*tag= b
FT mat_peptide 176..700
FT /*tag= c
XX
PN WO9905280-A1.
XX
PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-US15423.
XX
PR 25-JUL-1997; 97US-0900905.
XX
PA (SCHE) SCHERING CORP.
XX
PI Bazan JF;
XX
DR WPI; 1999-142935/12.
DR P-PSDB; W95003.
XX
PT Newly isolated or recombinant polynucleotide encoding mammalian
PT cytokine interleukin-B30 (IL-B30), including fragments - useful for
PT regulating activation, development, differentiation and function of
PT various cell types, and for diagnosing and treating conditions
PT associated with IL-B30
XX
PS Claim 3; Page 9-10; 83pp; English.
XX
CC This cDNA encodes a mouse cytokine interleukin-B30 (IL-B30) polypeptide.
CC Host cells containing a vector comprising the IL-B30 nucleic acid are
CC used for the recombinant production of the protein. The polynucleotides
CC are useful for diagnosis of IL-B30 mediated conditions, and forensic
CC science (e.g. to distinguish rodent from human, or as a marker to
CC distinguish between different cells exhibiting differential expression or
CC modification patterns). The IL-B30 (including fragments), together with
CC antibodies that bind to IL-B30 are useful for teaching purposes. They
CC are also used for treating conditions associated with abnormal physiology
CC or development, including inflammatory conditions. The polypeptide
CC cytokine should mediate cytokine synthesis and proliferation in cells.
CC IL-B30 is useful for drug screening to identify compounds having binding
CC affinity to IL-B30.
XX
SQ Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;

Query Match 68.6%; Score 391; DB 20; Length 1203;
Best Local Similarity 81.8%; Pred. No. 7.5e-105;
Matches 464; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

```
Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          |||||||   ||||||| |||||  |||| | |||| ||||||||    ||||||||
Db    113 atgctggattgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctg 172

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||| || | ||||| |||| |||| ||||||||||||||| || | ||| |||
Db    173 gctgtgcctaggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctc 232

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
          ||||  || |||||||  |||||| ||| || |||||| ||| ||   ||||||||||
Db    233 tgcatgctagcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaa 292

Qy    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
          |  ||||||||||||| |||| |||| |||| |||||||||| |||||| ||||||||
Db    293 gaggatgaagagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgaccca 352

Qy    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
          |||||||||| |||||||| |||||||||||||||||||||| ||| ||||||  |||| 
Db    353 caaggactcaaggacaacagccagttctgcttgcaaaggatccgccaaggtctggctttt 412

Qy    298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||| 
Db    413 tataagcacctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagc 472

Qy    358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417
          ||  ||| ||| ||||| |||||||||| || |||||||||||||| ||||| |||| | 
Db    473 cccatggagcaacttcacacctccctactaggactcagccaactcctccagccagaggat 532

Qy    418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477
          |||| | ||||||| || ||||| || ||||| ||| | || |||| ||||||||| | | 
Db    533 caccccgggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgcccc 592

Qy    478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537
          ||||||||| | ||| |||||||| |||||||||||||| |||| |||| |||||||||
Db    593 cttctccgttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtc 652

Qy    538 TTTGCCCATGGAGCAGCAACCCTGAGT 564
          |||||||| |||||||||| |||| |
Db    653 tttgcccacggagcagcaactctgact 679
```

RESULT 9
A52578
ID A52578 standard; cDNA; 1203 BP.
XX
AC A52578;
XX
DT 27-SEP-2000 (first entry)

```
XX
DE   Murine interleukin-B30 (IL-B30) cDNA.
XX
KW   Interleukin-B30; IL-B30; murine; mouse; cytokine; cellular signalling;
KW   immune response; haematopoietic cell; IL-6 homologue; GCSF homologue;
KW   granulocyte colony stimulating factor; immune disorder;
KW   inflammatory disease; autoimmune disease; antigen; antibody; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..703
FT                   /*tag= a
FT                   /product= "Murine IL-B30 precursor"
FT   sig_peptide     113..175
FT                   /*tag= b
FT   mat_peptide     176..700
FT                   /*tag= c
FT                   /product= "Mature murine IL-B30"
XX
PN   US6060284-A.
XX
PD   09-MAY-2000.
XX
PF   24-JUL-1998;   98US-0122443.
XX
PR   25-JUL-1997;   97US-0053765.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 2000-364420/31.
DR   P-PSDB; B01982.
XX
PT   Novel recombinant DNA encoding cytokines especially interleukin-B30
PT   useful as probes or primers for diagnosing immune disorders including
PT   autoimmune or chronic inflammatory conditions -
XX
PS   Claim 3; Column 7-10; 32pp; English.
XX
CC   This sequence represents cDNA encoding murine interleukin-B30 (IL-B30).
CC   IL-B30 is a novel cytokine, exhibiting significant homology to IL-6 and
CC   GCSF (granulocyte colony stimulating factor). Cytokines play a critical
CC   role in signalling between immune or other cells during an immune
CC   response. The precise role if IL-B30 is not yet known - it is likely to
CC   have either a stimulatory or an inhibitory effect on haematopoietic cells
CC   such as T-cells, B-cells, natural killer (NK) cells and macrophages.
CC   Alternatively, it may affect vascular physiology or development, or have
CC   neuronal effects. IL-B30, its fragments, IL-B30 nucleotides, agonists and
CC   antagonists are useful in the diagnosis and treatment of disorders
CC   associated with abnormal expression or activity of IL-B30 e.g.,
CC   inflammatory or autoimmune disorders. Nucleotides encoding IL-B30 are
CC   useful for recombinant expression of IL-B30 in a host cell, and as a
CC   source of probes and primers. The IL-B30 probes and primers can be used
CC   to detect levels of IL-B30 expression in samples from patients suspected
CC   of having an inflammatory or autoimmune disorder. IL-B30 nucleotides may
CC   also be used to identify homologous genes in other species. IL-B30
CC   protein or its fragments are useful as antigens for raising antibodies to
CC   various linear and conformational epitopes. Such antibodies may be used
CC   to detect levels of IL-B30 protein in a sample.
XX
SQ   Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;

  Query Match             68.6%;  Score 391;  DB 21;  Length 1203;
  Best Local Similarity   81.8%;  Pred. No. 7.5e-105;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
           |||||||   |||||||  |||||  ||||  |  ||||  ||||||||    ||||||||
Db     113 atgctggattgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctg 172

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
           |||||||||  ||  |  |||||  ||||  ||||  ||||||||||||||||  ||  |  |||  |||
Db     173 gctgtgcctaggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctc 232

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
           ||||  ||  |||||||  ||||||  |||  ||  ||||||  |||  ||   ||||||||||  .
Db     233 tgcatgctagcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaa 292

Qy     178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
           |  ||||||||||||||  ||||  ||||  ||||  ||||||||||  ||||||  ||||||||
Db     293 gaggatgaagagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgaccca 352

Qy     238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
           ||||||||||  |||||||||  ||||||||||||||||||||||||  |||  ||||||  ||||
Db     353 caaggactcaaggacaacagccagttctgcttgcaaaggatccgccaaggtctggctttt 412

Qy     298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357
           |||  ||  |  |||||  |   ||  ||  ||  ||||  ||||||||||  ||||  ||||||||||||
Db     413 tataagcacctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagc 472

Qy     358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417
           ||  |||  |||  |||||  ||||||||||  ||  ||||||||||||||  |||||  ||||  |
Db     473 cccatggagcaacttcacacctccctactaggactcagccaactcctccagccagaggat 532

Qy     418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477
           ||||  |  |||||||  ||  |||||  ||  |||  |  ||  ||||  ||||||||  |  |
Db     533 caccccccgggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgcccc 592

Qy     478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537
           ||||||||  |  |||  ||||||||  ||||||||||||||||  ||||  ||||  |||||||||
Db     593 cttctccgttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtc 652

Qy     538 TTTGCCCATGGAGCAGCAACCCTGAGT 564
           ||||||||  ||||||||||  ||||  |
Db     653 tttgcccacggagcagcaactctgact 679


RESULT  10
A42676
ID   A42676 standard; cDNA; 412 BP.
XX
AC   A42676;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1416.
XX
KW   Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021990-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24205.
XX
PR   15-OCT-1998;   98US-0104435.
XX
PA   (GEMY ) GENETICS INST INC.
```

```
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR   WPI; 2000-317937/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 459; 618pp; English.
XX
CC   A41261 to A43419 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include: chemotactic;
CC   proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;
CC   haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;
CC   antifungal; antiviral; antidiabetic; antiasthmatic; vulnerary; antiulcer;
CC   osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;
CC   cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be
CC   used for gene therapy and in vaccines. The sESTs are useful as probes for
CC   the identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression
CC   and psoriasis. A43420 to A43425 represent linker variants which are
CC   given in the exemplification of the present invention.
XX
SQ   Sequence 412 BP; 115 A; 101 C; 130 G; 66 T; 0 other;

  Query Match             42.6%;  Score 242.8;  DB 21;  Length 412;
  Best Local Similarity   99.2%;  Pred. No. 1.2e-61;
  Matches  244;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     164 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 223

Qy      61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     224 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 283

Qy     121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     284 tgcacactggcctggagtgcacatccactagtgggacacatggatctaagagaagaggga 343

Qy     181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     344 gatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgaccccca 403

Qy     241 GGACTC 246
          | ||||
Db     404 gaactc 409


RESULT  11
Z37263
ID   Z37263 standard; DNA; 2398 BP.
XX
AC   Z37263;
XX
DT   01-FEB-2000  (first entry)
XX
DE   SGRF coding sequence.
XX
KW   SGRF; human; Interleukin-6 G-CSF related factor; cell proliferation;
KW   immune system; haematopoietic system; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9954357-A1.
XX
PD   28-OCT-1999.
XX
PF   14-APR-1999;   99WO-JP01997.
XX
PR   14-APR-1998;   98JP-0121805.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Hirata Y;
XX
DR   WPI; 2000-013230/01.
DR   P-PSDB; Y54606.
XX
PT   Novel cytokine-like protein, with activity of supporting proliferation
PT   of myeloid cells, useful in treating abnormality of cell proliferation
PT   in immune and haematopoiesis systems  -
XX
PS   Disclosure; Fig 6; 69pp; Japanese.
XX
CC   This sequence encodes the Interleukin-6 G-CSF related factor (SGRF)
CC   protein of the invention. The protein is a member of the IL-6/G-CSF/MSF
CC   family. The protein can be used in drugs for treating diseases due to
CC   abnormality of cell proliferation in the immune system and haematopoietic
CC   system.
XX
SQ   Sequence 2398 BP; 599 A; 560 C; 633 G; 606 T; 0 other;

  Query Match             33.8%;  Score 192.8;  DB 21;  Length 2398;
  Best Local Similarity   74.3%;  Pred. No. 1.2e-46;
  Matches  309;  Conservative    0;  Mismatches    2;  Indels  105;  Gaps    1;

Qy     260 AGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGG 319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 agttctgcttgcaaaggatccaccagggtctgatttttttatgagaagctgctaggatcgg 1164

Qy     320 ATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCT 379
           |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db    1165 atattttcacaggggagccttctctgctccctgatagccctgtgggccagcttcatgcct 1224

Qy     380 CCCTACTGGGCCTCAGCCAACTCCTG---------------------------------- 405
           |||||||||||||||||||||||||
Db    1225 ccctactgggcctcagccaactcctgcaggtatgaagtaggggcgtggaggatgggggct 1284

Qy     406 ------------------------------------------------------------ 405

Db    1285 tgcaggtgtcagagacagagggttgggggttaagggtttagagtcttctctgactgtgtc 1344

Qy     406 -----------CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 454
                       ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 ctatgtcctttcagcctgagggtcaccactgggagactcagcagattccaagcctcagtc 1404

Qy     455 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 514
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 ccagccagccatggcagcgtctccttctccgcttcaaaatccttcgcagcctccaggcct 1464

Qy     515 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 ttgtggctgtagccgcccgggtctttgcccatggagcagcaaccctgagtccctaa 1520


RESULT  12
C66548/c
ID   C66548 standard; DNA; 121162 BP.
XX
AC   C66548;
```

XX
DT 19-FEB-2001 (first entry)
XX
DE Human kinesin-like protein HKLP coding sequence contig SEQ ID NO: 1.
XX
KW Human; kinesin-like protein; HKLP; KIF1; cell division; cancer;
KW intracellular transport; neurological disorder; infertility;
KW biallelic marker; spontaneous abortion; neonatal chromosome disorder;
KW aneuploidy; ds.
XX
OS Homo sapiens.
XX
PN WO200063375-A1.
XX
PD 26-OCT-2000.
XX
PF 20-APR-2000; 2000WO-IB00562.
XX
PR 20-APR-1999; 99US-0130217.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Dufaure-Gare I, Grel P;
XX
DR WPI; 2000-665242/64.
XX
PT An isolated or purified human kinesin-like protein (HKLP) encoding
PT polynucleotide used to detect HKLP polynucleotides in a sample
PT comprises a contiguous span of at least 12 nucleotides -
XX
PS Claim 1; Page 143-175; 199pp; English.
XX
CC The present invention describes the coding and protein sequences of the
CC human kinesin-like protein HKLP. It is thought that the protein could be
CC involved in neurological disorders, infertility, spontaneous abortion,
CC neonatal chromosome disorders, aneuploidy and cancers. This is due to its
CC function in the movement of microtubules. The protein shows homology to
CC the murine KIF1A and KIF1B proteins. The sequences disclosed in the
CC invention can be used in the isolation of similar human proteins and in
CC vector production. In addition, the biallelic markers shown can be used
CC in disease diagnosis and population studies.
XX
SQ Sequence 121162 BP; 33272 A; 24108 C; 25842 G; 37919 T; 21 other;

Query Match 6.7%; Score 38; DB 21; Length 121162;
Best Local Similarity 54.2%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 190 ACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGG 249
Db 87799 AATAGCAATGATTTCACTGTAATAATGTATTTATATGCCTCTCTCCCCCAGACAGCAAGC 87740

Qy 250 GACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTG 309
Db 87739 GTTTTCCCACAGGTCTCGACACACAGAAGATACTCTGGAAAGATTTGTTAAAACAATGGG 87680

Qy 310 CTAGGATCGGATATTTTCACAG 331
Db 87679 CTATATTTAGTTATTTTCACAG 87658

RESULT 13
Z55567
ID Z55567 standard; cDNA; 567 BP.
XX
AC Z55567;
XX
DT 14-MAR-2000 (first entry)
XX
DE Feline interferon-alpha (IFN-alpha) clone #2 cDNA.
XX
KW Interferon-alpha; IFN-alpha; antibody; feline; inhibitor;
KW immune response; immunoregulation; tumour; cancer; autoimmune disease;
KW vaccine; ss.
XX
OS Felis catus.
XX
PN WO9961618-A2.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US11942.
XX
PR 29-MAY-1998; 98US-0087306.
XX
PA (HESK-) HESKA CORP.
XX
PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
XX
DR WPI; 2000-072623/06.
DR P-PSDB; Y58225.
XX
PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
PT useful for treating or preventing e.g. tumors or autoimmune disease -
XX
PS Claim 1j; Page 241-242; 264pp; English.
XX
CC Sequences Z55567-Z55568, Z55571-Z55572, Z55569-Z55570 and
CC Z55573-Z55574 represent cDNA sequences encoding feline encoded
CC and mature interferon-alpha (IFN-alpha) clones #2
CC and #3 respectively. The invention relates to canine
CC interleukin-4 (IL-4), canine or feline Flt-3 ligand, canine
CC or feline CD40, canine or feline CD154 (CD40 ligand), canine IL-5,
CC canine IL-13, feline IFN-alpha and feline granulocyte macrophage
CC colony-stimulating factor (GMCSF), and nucleotides which encode these
CC immunoregulatory proteins. The proteins, their associated
CC nucleic acids, specific antibodies and inhibitors may be used as
CC vaccines for therapeutic or prophylactic regulation of an immune
CC response in animals (particularly cats, dogs, horses and humans).
CC They may be used to treat autoimmune or infectious diseases including
CC allergies, tumours, inflammation and graft rejection, and to increase
CC the response from a co-administered antigen. The nucleotide sequences
CC can also be used for the recombinant production of a protein, while
CC nucleotide fragments are useful as probes, as amplification primers and
CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
CC The proteins may be used to raise antibodies and to screen for
CC modulators of activity, while the antibodies may be used in detection,
CC and in drug targetting.
XX
SQ Sequence 567 BP; 126 A; 167 C; 156 G; 118 T; 0 other;

Query Match 6.1%; Score 35; DB 21; Length 567;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 71; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 32 TGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCTGCCTGGA 91
Db 2 tggcgctgccctcttccttcttggtggccctggtggcgctgggctgcaactccgtctgct 61

Qy 92 CTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACATCCACTAG 151
Db 62 ctctgggctgtgacctgcctcagacccacggcctgctgaacaggagggccttgacgctcc 121

Qy 152 TGGGACACATG 162
Db 122 tgggacaaatg 132

RESULT 14
Z55568/c
ID Z55568 standard; cDNA; 567 BP.
XX
AC Z55568;
XX

```
DT   14-MAR-2000  (first entry)
XX
DE   Feline interferon-alpha (IFN-alpha) clone #2 cDNA complement.
XX
KW   Interferon-alpha; IFN-alpha; antibody; feline; inhibitor;
KW   immune response; immunoregulation; tumour; cancer; autoimmune disease;
KW   vaccine; ss.
XX
OS   Felis catus.
XX
PN   WO9961618-A2.
XX
PD   02-DEC-1999.
XX
PF   28-MAY-1999;   99WO-US11942.
XX
PR   29-MAY-1998;   98US-0087306.
XX
PA   (HESK-) HESKA CORP.
XX
PI   Sim G,  Yang S,  Dreitz MJ,  Wonderling RS;
XX
DR   WPI; 2000-072623/06.
DR   P-PSDB; Y58225.
XX
PT   Nucleic acids encoding immunoregulatory proteins from cats or dogs,
PT   useful for treating or preventing e.g. tumors or autoimmune disease -
XX
PS   Claim 1j; Page 243; 264pp; English.
XX
CC   Sequences Z55567-Z55568, Z55571-Z55572, Z55569-Z55570 and
CC   Z55573-Z55574 represent cDNA sequences encoding feline encoded
CC   and mature interferon-alpha (IFN-alpha) clones #2
CC   and #3 respectively. The invention relates to canine
CC   interleukin-4 (IL-4), canine or feline Flt-3 ligand, canine
CC   or feline CD40, canine or feline CD154 (CD40 ligand), canine IL-5,
CC   canine IL-13, feline IFN-alpha and feline granulocyte macrophage
CC   colony-stimulating factor (GMCSF), and nucleotides which encode these
CC   immunoregulatory proteins. The proteins, their associated
CC   nucleic acids, specific antibodies and inhibitors may be used as
CC   vaccines for therapeutic or prophylactic regulation of an immune
CC   response in animals (particularly cats, dogs, horses and humans).
CC   They may be used to treat autoimmune or infectious diseases including
CC   allergies, tumours, inflammation and graft rejection,  and to increase
CC   the response from a co-administered antigen. The nucleotide sequences
CC   can also be used for the recombinant production of a protein, while
CC   nucleotide fragments are useful as probes, as amplification primers and
CC   as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
CC   The proteins may be used to raise antibodies and to screen for
CC   modulators of activity, while the antibodies may be used in detection,
CC   and in drug targetting.
XX
SQ   Sequence 567 BP; 118 A; 156 C; 167 G; 126 T; 0 other;


  Query Match             6.1%;  Score 35;  DB 21;  Length 567;
  Best Local Similarity   54.2%;  Pred. No. 1;
  Matches   71;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy      32 TGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCTGCCTGGA 91
           ||  ||||||||   |      ||        ||| |   |||| ||| | | | |||
Db     566 TGGCGCTGCCCTCTTCCTTCTTGGTGGCCCTGGTGGCGCTGGGCTGCAACTCCGTCTGCT 507

Qy      92 CTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACATCCACTAG 151
           ||| | ||    | ||  | ||||  | | ||   |||  | |||| ||     | ||
Db     506 CTCTGGGCTGTGACCTGCCTCAGACCCACGGCCTGCTGAACAGGAGGGCCTTGACGCTCC 447

Qy     152 TGGGACACATG 162
           ||||||| |||
Db     446 TGGGACAAATG 436


RESULT  15
```

```
Z55569
ID   Z55569 standard; cDNA; 567 BP.
XX
AC   Z55569;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Feline interferon-alpha (IFN-alpha) clone #3 cDNA.
XX
KW   Interferon-alpha; IFN-alpha; antibody; feline; inhibitor;
KW   immune response; immunoregulation; tumour; cancer; autoimmune disease;
KW   vaccine; ss.
XX
OS   Felis catus.
XX
PN   WO9961618-A2.
XX
PD   02-DEC-1999.
XX
PF   28-MAY-1999;   99WO-US11942.
XX
PR   29-MAY-1998;   98US-0087306.
XX
PA   (HESK-) HESKA CORP.
XX
PI   Sim G,  Yang S,  Dreitz MJ,  Wonderling RS;
XX
DR   WPI; 2000-072623/06.
DR   P-PSDB; Y58226.
XX
PT   Nucleic acids encoding immunoregulatory proteins from cats or dogs,
PT   useful for treating or preventing e.g. tumors or autoimmune disease -
XX
PS   Claim 1j; Page 244-245; 264pp; English.
XX
CC   Sequences Z55567-Z55568, Z55571-Z55572, Z55569-Z55570 and
CC   Z55573-Z55574 represent cDNA sequences encoding feline encoded
CC   and mature interferon-alpha (IFN-alpha) clones #2
CC   and #3 respectively. The invention relates to canine
CC   interleukin-4 (IL-4), canine or feline Flt-3 ligand, canine
CC   or feline CD40, canine or feline CD154 (CD40 ligand), canine IL-5,
CC   canine IL-13, feline IFN-alpha and feline granulocyte macrophage
CC   colony-stimulating factor (GMCSF), and nucleotides which encode these
CC   immunoregulatory proteins. The proteins, their associated
CC   nucleic acids, specific antibodies and inhibitors may be used as
CC   vaccines for therapeutic or prophylactic regulation of an immune
CC   response in animals (particularly cats, dogs, horses and humans).
CC   They may be used to treat autoimmune or infectious diseases including
CC   allergies, tumours, inflammation and graft rejection,  and to increase
CC   the response from a co-administered antigen. The nucleotide sequences
CC   can also be used for the recombinant production of a protein, while
CC   nucleotide fragments are useful as probes, as amplification primers and
CC   as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
CC   The proteins may be used to raise antibodies and to screen for
CC   modulators of activity, while the antibodies may be used in detection,
CC   and in drug targetting.
XX
SQ   Sequence 567 BP; 123 A; 169 C; 158 G; 117 T; 0 other;


  Query Match             6.1%;  Score 35;  DB 21;  Length 567;
  Best Local Similarity   54.2%;  Pred. No. 1;
  Matches   71;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy      32 TGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCTGCCTGGA 91
           ||  ||||||||   |      ||        ||| |   |||| ||| | | | |||
Db      2 tggcgctgccctcttccttcttggtggccctggtggcgctgggctgcaactccgtctgct 61

Qy      92 CTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACATCCACTAG 151
           ||| | ||    | ||  | ||||  | | ||   |||  | |||| ||     | ||
Db     62 ctctgggctgtgacctgcctcagacccacggcctgctgaacaggagggccttgacgctcc 121

Qy     152 TGGGACACATG 162
```

```
          ||||||| |||
Db    122 tgggacaaatg 132


Search completed: June 21, 2001, 02:08:54
Job time: 4583 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2001, 23:16:50 ; Search time 1723.79 Seconds
(without alignments)
3125.744 Million cell updates/sec

Title: US-09-558-474-1
Perfect score: 570
Sequence: 1 ATGCTGGGGAGCAGAGCTGT.....CAGCAACCCTGAGTCCCTAA 570

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	404.8	71.0	833	104	AJ277049	AJ277049 AJ277049
2	393	68.9	492	7	AA418955	AA418955 zw01c10.r
3	321.2	56.4	1150	137	BE572471	BE572471 601331259
c 4	232.8	40.8	542	102	AI796983	AI796983 we25e07.x
5	231.4	40.6	897	137	BE571481	BE571481 601333907
6	220	38.6	372	155	C06368	C06368 C06368 Huma
7	210.4	36.9	430	141	BE847383	BE847383 uw26b04.y
8	206.8	36.3	719	153	BG435629	BG435629 602506911
c 9	183	32.1	496	175	BG236314	BG236314 naf26b02.
c 10	150	26.3	457	117	AW511070	AW511070 hd38e10.x
c 11	140.2	24.6	367	145	BF189762	BF189762 235459 MA
c 12	138.6	24.3	575	31	AV598349	AV598349 AV598349
c 13	122.8	21.5	399	114	AW316433	AW316433 14841 MAR
c 14	113.2	19.9	566	165	BE237060	BE237060 146167 MA
15	106.4	18.7	323	129	BB270313	BB270313 BB270313
16	102	17.9	320	130	BB274745	BB274745 BB274745
17	91.6	16.1	137	188	T18577	T18577 hbc2125 Hum
c 18	90	15.8	400	14	AI004049	AI004049 ou11d12.x
c 19	58	10.2	369	111	AW072323	AW072323 wz99h07.x
c 20	54	9.5	397	7	AA418747	AA418747 zw01c10.s
c 21	39.4	6.9	404	170	BF834812	BF834812 CM3-HT096
22	38.8	6.8	579	141	BE884343	BE884343 601505619
23	38.6	6.8	1101	219	CNS00LXJ	AL078875 Drosophil
c 24	38.4	6.7	198	11	AA738190	AA738190 nx12f02.s
c 25	38	6.7	558	140	BE809252	BE809252 214651 MA
c 26	37.8	6.6	459	234	AQ852080	AQ852080 LMAJFV1_1
27	37.4	6.6	889	145	BF204344	BF204344 601866418
c 28	37.2	6.5	383	7	AA415467	AA415467 vc99g05.s
c 29	37	6.5	421	12	AA794290	AA794290 vu67g02.r
30	37	6.5	539	220	CNS0276O	AL184281 Tetraodon
31	37	6.5	600	156	C77322	C77322 C77322 Mous
c 32	36.8	6.5	355	142	BE935507	BE935507 MR2-MT012
c 33	36.8	6.5	440	7	AA434734	AA434734 ve17f02.r
c 34	36.8	6.5	522	136	BE494337	BE494337 WHE1253_C
35	36.8	6.5	831	220	CNS01YR4	AL173353 Tetraodon
c 36	36.6	6.4	376	2	AA110023	AA110023 ml60h02.r
c 37	36.4	6.4	609	175	BG300864	BG300864 HVSMEb001
c 38	36.4	6.4	646	143	BF043357	BF043357 BP250017A
c 39	36.2	6.4	409	138	BE656580	BE656580 UI-M-BH0-
c 40	36	6.3	324	167	BE400665	BE400665 AWB006.D0
41	36	6.3	469	121	AW839214	AW839214 CM0-LT006
42	36	6.3	933	106	AL578453	AL578453 AL578453
43	36	6.3	989	106	AL578994	AL578994 AL578994
44	36	6.3	1037	106	AL574127	AL574127 AL574127
c 45	35.8	6.3	469	16	AI120671	AI120671 ub71d01.r

ALIGNMENTS

```
RESULT   1
AJ277049
LOCUS       AJ277049      833 bp    mRNA            EST       30-MAR-2000
DEFINITION  AJ277049 Placenta Matchmaker cDNA library(Clontech) Homo sapiens
            cDNA clone BHA2; mRNA sequence.
ACCESSION   AJ277049
VERSION     AJ277049.1  GI:7342435
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   Bach,S., Portetelle,D. and Vandenbol,M.
  TITLE     Isolation of new human EST sequences by using the MSB4 yeast gene
            as bait in the two-hybrid screening of a Human Placenta cDNA
            library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Bach SL
            Microbiology Unit
            Gembloux Agricultural University
            Avenue Marechal Juin 6, B-5030 Gembloux, Belgium.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BHA2"
                     /clone_lib="Placenta Matchmaker cDNA library(Clontech)"
                     /sex="Female"
                     /tissue_type="Placenta"
BASE COUNT      227 a    197 c    200 g    209 t
ORIGIN


  Query Match             71.0%;  Score 404.8;  DB 104;  Length 833;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-97;
  Matches  406;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     163 GATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGA 222
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 GATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGA 156

Qy     223 GATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCAC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157 GATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCAC 216

Qy     283 CAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCT 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 CAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCT 276

Qy     343 CTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTC 402
           |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db     277 CTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTC 336

Qy     403 CTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAG 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 CTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAG 396

Qy     463 CCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCT 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 CCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCT 456

Qy     523 GTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 GTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 504


RESULT   2
AA418955
LOCUS       AA418955      492 bp    mRNA            EST       16-OCT-1997
DEFINITION  zw01c10.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:768018
            5', mRNA sequence.
ACCESSION   AA418955
VERSION     AA418955.1  GI:2080765
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:768018"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      117 a    136 c    136 g    103 t
ORIGIN


  Query Match             68.9%;  Score 393;  DB 7;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-94;
  Matches  393;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     163 GATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGA 222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     100 GATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGA 159

Qy     223 GATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCAC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     160 GATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCAC 219

Qy     283 CAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCT 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     220 CAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCT 279

Qy     343 CTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTC 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     280 CTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTC 339

Qy     403 CTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAG 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     340 CTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAG 399

Qy     463 CCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCT 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     400 CCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCT 459

Qy     523 GTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCA 555
           |||||||||||||||||||||||||||||||||
Db     460 GTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCA 492


RESULT   3
BE572471
LOCUS       BE572471     1150 bp    mRNA            EST       15-AUG-2000
DEFINITION  601331259F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3708687 5',
            mRNA sequence.
ACCESSION   BE572471
VERSION     BE572471.1  GI:9816191
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1150)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8928  row: f  column: 16
            High quality sequence stop: 716.
FEATURES             Location/Qualifiers
     source          1. .1150
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3708687"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      330 a    259 c    329 g    231 t      1 others
ORIGIN


  Query Match             56.4%;  Score 321.2;  DB 137;  Length 1150;
  Best Local Similarity   79.4%;  Pred. No. 3e-75;
  Matches  456;  Conservative    0;  Mismatches  108;  Indels   10;  Gaps    6;

Qy       1 ATGCTGGGGAGCAGAGCTGTAATGCTGCT-GTTGCTGCTGCCCTGGACAGCTCAGGGCAG 59
           |||||||  |||||||  |||||  ||||  ||  ||||  ||||||||    ||||||||  |
Db      78 ATGCTGGATTGCAGAGCAGTAATAATGCTAGTGGCTGTTGCCCTGGGTCACTCAGGGCCG 137

Qy      60 AG--CTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAG 117
            |        | |  ||  |  |||||  ||||  ||||  ||||||||||||||||||  ||  |  |||
Db     138 TGGCTGTGCCGTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAAT 197

Qy     118 CTCTGC--ACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAG 172
           ||||||      |  |||||||  ||||||  |||  ||  ||||||  |||  ||    |||||||
Db     198 CTCTGCGATGCGTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAG 257

Qy     173 AAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTG 232
           ||||  |  ||||||||||||||  ||||  ||||  ||||  ||||||||||  ||||||  ||||
Db     258 AAGAAGAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTG 317

Qy     233 ACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGA 292
           ||||  ||||||||||  |||||||||  ||||||||||||||||||||||||  |||  ||||||
Db     318 ACCCACAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGG 377

Qy     293 TTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCC-T 351
           |||||||  ||  |  |||||  |  ||  ||  ||  ||||  ||||||||||  ||||  |||||  |
Db     378 CTTTTTATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCGT 437

Qy     352 GATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCT 411
           ||||||||  |||  |||  |||||  ||||||||||  ||  ||||||||||||||  |||||
Db     438 GATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCA 497

Qy     412 GAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCC-TCAGTCCCAGCCAGCCATGGCA 470
           ||||  |||||  |  |||||||  ||  |||||  ||  ||||  |  |||  |  ||  ||||  |||||
Db     498 GAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTTGAGTTCTAGTCAGCAGTGGCA 557

Qy     471 GCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGC 530
           |||  |  ||||||||||  |  |||  ||||||||  ||||||||||||||||  ||||  ||||  ||
```

```
Db    558 GCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGC 617

Qy    531 CCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGT 564
          ||||||||| | ||| |||||||||| |||| |
Db    618 CCGGGTCTTGGGCCACGGAGCAGCAACTCTGACT 651


RESULT   4
AI796983/c
LOCUS       AI796983      542 bp    mRNA            EST       18-DEC-1999
DEFINITION  we25e07.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2342148 3',
            mRNA sequence.
ACCESSION   AI796983
VERSION     AI796983.1  GI:5362455
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 868   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2342148"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT      150 a    125 c    140 g    127 t
ORIGIN


  Query Match             40.8%;  Score 232.8;  DB 102;  Length 542;
  Best Local Similarity   99.2%;  Pred. No. 1e-51;
  Matches  234;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    335 AGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCA 394
          ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    542 AGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCA 483

Qy    395 GCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 GCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 423

Qy    455 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 363

Qy    515 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 307


RESULT   5
BE571481
LOCUS       BE571481      897 bp    mRNA            EST       15-AUG-2000
DEFINITION  601333907F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3711529 5',
            mRNA sequence.
ACCESSION   BE571481
VERSION     BE571481.1  GI:9815201
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8935  row: m  column: 02
            High quality sequence stop: 597.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3711529"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      239 a    201 c    262 g    194 t      1 others
ORIGIN


  Query Match             40.6%;  Score 231.4;  DB 137;  Length 897;
  Best Local Similarity   73.4%;  Pred. No. 2.8e-51;
  Matches  391;  Conservative    0;  Mismatches  131;  Indels   11;  Gaps    7;

Qy      7 GGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTG 66
          ||  ||||||| |||||  |||| | |||| ||||||||    ||||||||   | ||||
Db    117 GGATGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGC-CTGGTGTG 175

Qy     67 CCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACA 126
          ||| || | ||||| |||| ||| ||||||||||||||| || | ||| |||||||
Db    176 CCTAGGAGTAGCAGTCCTGAGTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATG 235

Qy    127 CTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAGGGAGAT 183
          || |||||||  |||||| ||| || |||||| ||| ||   |||||||||||| |  |||
Db    236 CTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGAT 295

Qy    184 GAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGA 243
          |||||||||| |||| |||| |||| |||||||||| |||||| |||||||| ||||||
Db    296 GAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGA 355

Qy    244 CTCAGGGACAACAGTCAGTTCTGCTTGC-AAAGGATCCACCAGGGTCTGATTTTTTATGA 302
```

```
          |||| |||||||||| |||||||||||| ||||||||| ||| ||||||  ||||||| |
Db    356 CTCAAGGACAACAGCCAGTTCTGCTTGCAAAAGGATCCGCCAAGGTCTGGCTTTTTATAA 415

Qy    303 GAAGCTGCTAGGATCGGATATTTTC-ACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTG 361
          | | ||||| |  || || || ||| | ||||||||| | ||| |||||||||   ||
Db    416 GCACCTGCTTGACTCTGACATCTTCAAGAGGGGAGCCGTGTCTACTCCCTGATTAGCCCC 475

Qy    362 TGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACC 421
            |      ||||  |||||||||| || |||||||||   |  |   |    ||    |
Db    476 ATGGAGCAATTCACACCTCCCTACTAGGACTCAGCCAAACTCCTCCAGCCAGAGGATCAC 535

Qy    422 ACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTC 481
            |  |||||| || ||||| || || || ||| | || ||||  |||       | |||||
Db    536 CCCCGGAGACCCAACAGATGCCCAG-CTGAGTTCTAGTCAGCAGTGG--AGCGCCCCTTC 592

Qy    482 TCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGG 534
          |||| | ||| |||  |   |  ||||||| |||||||    |||| ||| ||
Db    593 TCCGTTCCAAGATC--TTCGAAGCTCCAGGGCTTTGTGCGCATAGCTGCCGGG 643


RESULT   6
C06368
LOCUS       C06368        372 bp    mRNA            EST       16-OCT-1996
DEFINITION  C06368 Human pancreatic islet Homo sapiens cDNA clone hbc5699, mRNA
            sequence.
ACCESSION   C06368
VERSION     C06368.1  GI:1503144
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   Takeda,J.
  TITLE     Human pancreatic islet ESTs
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Jun Takeda
            Institute for Molecular and Cellular Regulation, Gunma University
            3-39-15 Showa-machi, Maebashi Gunma 371, Japan
            Tel: 272-20-8856
            Fax: 272-20-8896
            Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hbc5699"
                     /clone_lib="Human pancreatic islet"
                     /note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
                     I; mRNA was prepared from normal adult human islets. cDNA
                     was directionally synthesized from the Xho I in the vector
                     to the EcoRI site. cDNA was size fractionated to remove
                     sequences <1000 bp in size."
BASE COUNT      107 a     88 c    116 g     59 t      2 others
ORIGIN


  Query Match           38.6%;  Score 220;  DB 155;  Length 372;
  Best Local Similarity   99.1%;  Pred. No. 2.5e-48;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 210

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 270

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGAGGGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    271 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGATCTAAGAGAAGANGGA 330
```

```
Qy    181 GATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGA 222
          |||||||||||||| ||||||||||||||||||||||||||||
Db    331 GATGAAGAGACTACNAATGATGTTCCCCATATCCAGTGTGGA 372


RESULT   7
BE847383
LOCUS       BE847383      430 bp    mRNA            EST       26-SEP-2000
DEFINITION  uw26b04.y1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3417775 5', mRNA sequence.
ACCESSION   BE847383
VERSION     BE847383.1  GI:10305722
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1093587
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3417775"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      120 a     99 c    123 g     88 t
ORIGIN


  Query Match           36.9%;  Score 210.4;  DB 141;  Length 430;
  Best Local Similarity   81.3%;  Pred. No. 9.5e-46;
  Matches  257;  Conservative    0;  Mismatches   56;  Indels    3;  Gaps    1;

Qy      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60
          |||||||   |||||| |||||  |||| | |||| ||||||||    ||||||||
Db    111 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 170

Qy     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120
          |||||||| || | ||||| |||| |||| ||||||||||||||| || | ||| |||
Db    171 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 230

Qy    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177
          ||||  || |||||||  |||||| ||| || |||||| ||| ||   ||||||||||
Db    231 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 290

Qy    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237
          |  |||||||||||| |||| |||| |||| |||||||||| |||||| ||||||||
Db    291 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 350
```

```
Qy    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297
          ||||||||| |||||||| |||||||||||||||||||| ||| ||||||  ||||
Db    351 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 410

Qy    298 TATGAGAAGCTGCTAG 313
          ||| || | ||||| |
Db    411 TATAAGCACCTGCTTG 426


RESULT   8
BG435629
LOCUS       BG435629      719 bp    mRNA            EST       14-MAR-2001
DEFINITION  602506911F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4604139 5',
            mRNA sequence.
ACCESSION   BG435629
VERSION     BG435629.1  GI:13342135
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1339  row: e  column: 04
            High quality sequence stop: 496.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4604139"
                     /clone_lib="NIH_MGC_79"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      185 a    184 c    179 g    171 t
ORIGIN


  Query Match             36.3%;  Score 206.8;  DB 153;  Length 719;
  Best Local Similarity   99.0%;  Pred. No. 1e-44;
  Matches  208;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    361 GTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 420
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCAC 60

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 120

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 180

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    181 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 210


RESULT   9
BG236314/c
LOCUS       BG236314      496 bp    mRNA            EST       12-FEB-2001
DEFINITION  naf26b02.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:4142138 3',
            mRNA sequence.
ACCESSION   BG236314
VERSION     BG236314.1  GI:12750161
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: M. Bento Soares, Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D. and M.Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .496
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4142138"
                     /clone_lib="Soares_NPBMC"
                     /tissue_type="lymphocyte"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: blood; Vector: pT7T3D-Pac; Site_1: NotI;
                     Site_2: EcoRI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCGGGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested  with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3  vector. Library
                     is normalized; constructed in the laboratory of M. Bento
                     Soares (University of Iowa)."
BASE COUNT      134 a    114 c    119 g    129 t
ORIGIN


  Query Match             32.1%;  Score 183;  DB 175;  Length 496;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-38;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    388 GGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGC 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 GGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGC 437

Qy    448 CTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTC 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 CTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTC 377

Qy    508 CAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCC 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 CAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCC 317

Qy    568 TAA 570
```

```
                ||||
Db    316 TAA 314


RESULT 10
AW511070/c
LOCUS       AW511070      457 bp    mRNA            EST       03-MAR-2000
DEFINITION  hd38e10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2911818 3', mRNA sequence.
ACCESSION   AW511070
VERSION     AW511070.1  GI:7149148
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2911818"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      129 a    105 c    106 g    117 t
ORIGIN


  Query Match             26.3%;  Score 150;  DB 117;  Length 457;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-29;
  Matches  150;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    421 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 CACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTT 398

Qy    481 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397 CTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTT 338

Qy    541 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          ||||||||||||||||||||||||||||||
Db    337 GCCCATGGAGCAGCAACCCTGAGTCCCTAA 308


RESULT 11
BF189762/c
LOCUS       BF189762      367 bp    mRNA            EST       02-NOV-2000
DEFINITION  235459 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BF189762
VERSION     BF189762.1  GI:11073131
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 58  row: E  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       75 a    100 c    107 g     85 t
ORIGIN


  Query Match             24.6%;  Score 140.2;  DB 145;  Length 367;
  Best Local Similarity   89.3%;  Pred. No. 5.3e-27;
  Matches  151;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    402 CCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCA 461
          | | ||||| ||||||||||||||||||||||| ||||||  ||||||| ||||||||||||
Db    270 CTTTCAGCCCGAGGGTCACCACTGGGAGACTGAGCAGACGCCAAGCCCCAGTCCCAGCCA 211

Qy    462 GCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGC 521
          ||| ||||| || ||||||||||| |||| ||||||||||||||||||||||||||||||
Db    210 GCCCTGGCAACGCCTCCTTCTCCGCCTCAAGATCCTTCGCAGCCTCCAGGCCTTTGTGGC 151

Qy    522 TGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
          |||||| |||||||||||| |||||||||||||||||| ||||| |  |||
Db    150 TGTAGCTGCCCGGGTCTTCGCCCATGGAGCAGCAACTCTGAGCCAGTAA 102


RESULT 12
AV598349/c
LOCUS       AV598349      575 bp    mRNA            EST       29-AUG-2000
DEFINITION  AV598349 Bos taurus cartilage fetus Bos taurus cDNA clone
            E1CA037E02 5', mRNA sequence.
ACCESSION   AV598349
VERSION     AV598349.1  GI:9715927
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 575)
  AUTHORS   Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
```

```
  TITLE     bovine cDNA sequencing
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .575
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1CA037E02"
                     /clone_lib="Bos taurus cartilage fetus"
                     /tissue_type="cartilage"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                     was deleted from a Not1 site"
BASE COUNT      140 a    156 c    168 g    111 t
ORIGIN


  Query Match             24.3%;  Score 138.6;  DB 31;  Length 575;
  Best Local Similarity   88.8%;  Pred. No. 1.6e-26;
  Matches  150;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     402 CCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCA 461
           | | |||||  ||||||||||||||||| ||| |||||| ||||||||  | |||||||||
Db     220 CTTTCAGCCCAAGGGTCACCACTGGGAAACTGAGCAGACTCCAAGCCCTATTCCCAGCCA 161

Qy     462 GCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGC 521
           ||||||||||||  ||||||||||  |||| |||||||| ||||||||||||||||||||
Db     160 GCCATGGCAGCGCCTCCTTCTCCGTCTCAAGATCCTTCGAAGCCTCCAGGCCTTTGTGGC 101

Qy     522 TGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAA 570
           |||||| |||||||||||||||| |||||||||| ||||| ||||||
Db     100 TGTAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGAGCCCCTAA 52


RESULT  13
AW316433/c
LOCUS       AW316433      399 bp    mRNA            EST       09-JUL-2000
DEFINITION  14841 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW316433
VERSION     AW316433.1  GI:6745689
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 11  row: O  column: 23
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT      100 a    114 c    101 g     84 t
ORIGIN


  Query Match             21.5%;  Score 122.8;  DB 114;  Length 399;
  Best Local Similarity   88.7%;  Pred. No. 2.5e-22;
  Matches  133;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy     259 CAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCG 318
           |||| |||||||||||||||||||||| || ||| |||||||||||||||||||| || ||
Db     289 CAGTCCTGCTTGCAAAGGATCCACCAAGGCCTGGTTTTTTATGAGAAGCTGCTGGGCTCA 230

Qy     319 GATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCC 378
           || |||||||||||||||||||||| | ||||||| || ||||||  |||||||| |||
Db     229 GACATTTTCACAGGGGAGCCTTCTCTACACCCTGATGGCTCTGTGGGCCAGCTTCACGCC 170

Qy     379 TCCCTACTGGGCCTCAGCCAACTCCTGCAG 408
           |||||||||||||||| |||||| |||||
Db     169 TCCCTACTGGGCCTCAGGCAACTCTTGCAG 140


RESULT  14
BE237060/c
LOCUS       BE237060      566 bp    mRNA            EST       10-JUL-2000
DEFINITION  146167 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE237060
VERSION     BE237060.1  GI:9021778
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
            Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
            ,W.W. and Keele,J.W.
  TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 46  row: N  column: 3
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 4BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
```

```
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from day 20 and day 40
                     embryos."
BASE COUNT      157 a    145 c    138 g    126 t
ORIGIN

  Query Match             19.9%;  Score 113.2;  DB 165;  Length 566;
  Best Local Similarity   84.7%;  Pred. No. 1e-19;
  Matches  127;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     259 CAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCG 318
           |||| |||||||||||| || || |  || ||| ||||||| |||||||||| || ||
Db     250 CAGTCCTGCTTGCAAAGAATTCATCGAGGCCTGGTTTTTTACGAGAAGCTGCTGGGCTCA 191

Qy     319 GATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCC 378
           ||||||||||||||||||||||||| |||||  || |||||||||  |||||||| |||
Db     190 GATATTTTCACAGGGGAGCCTTCTCTACTCCCAAATGGCCCTGTGGACCAGCTTCACGCC 131

Qy     379 TCCCTACTGGGCCTCAGCCAACTCCTGCAG 408
           ||| ||||||||||||  ||||| |||||
Db     130 TCCATACTGGGCCTCAGGGAACTCTTGCAG 101


RESULT  15
BB270313
LOCUS       BB270313      323 bp    mRNA            EST       07-JUL-2000
DEFINITION  BB270313 RIKEN full-length enriched, 10 days neonate cortex Mus
            musculus cDNA clone A830047G15 3', mRNA sequence.
ACCESSION   BB270313
VERSION     BB270313.1  GI:8966785
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .323
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A830047G15"
                     /clone_lib="RIKEN full-length enriched, 10 days neonate
                     cortex"
                     /tissue_type="cortex"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer {5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT       73 a     99 c     66 g     85 t
ORIGIN

  Query Match             18.7%;  Score 106.4;  DB 129;  Length 323;
  Best Local Similarity   78.0%;  Pred. No. 5.6e-18;
  Matches  128;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy     401 TCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCC 460
           |||  ||||| |||| ||||| | ||||||| || ||||| || | ||| | |   ||
Db       2 TCCCACAGCCAGAGGTTCACCCCCGGGAGACCCAACAGATGCCCAACCTGAATTTTAGTT 61

Qy     461 AGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGG 520
           ||| |||||||| | |||| ||| | ||| ||||||| ||||||||| ||||| |||
Db      62 AGCAATGGCAGCGCCCCCTTTTCCCTTCCAAGTTCCTTCGAAGCCTCCAGCCCTTTTTGG 121

Qy     521 CTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGT 564
           |  |||| |||||||||||||||| ||||||||||| |||| |
Db     122 CCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACT 165


Search completed: June 21, 2001, 01:24:55
Job time: 7685 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2001, 02:07:38 ; Search time 2434.93 Seconds
(without alignments)
7641.987 Million cell updates/sec

Title: US-09-558-474-3
Perfect score: 1203
Sequence: 1 CGCTTAGAAGTCGGACTACA.....TACCAATACGAACTGACAAA 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1203	100.0	1203	10	AX097623	AX097623 Sequence
	2	1203	100.0	1358	94	AF301619	AF301619 Mus muscu
	3	441.4	36.7	1025	89	AF301620	AF301620 Homo sapi
	4	439.8	36.6	1026	85	AB030000	AB030000 Homo sapi
	5	439.8	36.6	1055	9	AX048200	AX048200 Sequence
	6	391	32.5	570	10	AX097621	AX097621 Sequence
	7	151.6	12.6	4039	85	AB030001	AB030001 Homo sapi
c	8	150.6	12.5	155929	69	AC025574	AC025574 Homo sapi

	9	123.2	10.2	2300	7	AB030002 AB030002 Sus scrof
c	10	123.2	10.2	4649	7	AB004061 AB004061 domestic
c	11	62	5.2	7218	10	I66494 I66494 Sequence 14
c	12	43.8	3.6	90558	14	NCB24M22 AL390354 Neurospor
c	13	43.8	3.6	258622	66	AC020877 AC020877 Mus muscu
c	14	43.4	3.6	62336	77	AC087468 AC087468 Homo sapi
c	15	41.8	3.5	1111	53	CNS0763N AL430905 clone XBA
c	16	41	3.4	209826	73	AC068497 AC068497 Mus muscu
c	17	40.6	3.4	151924	88	AC068722 AC068722 Homo sapi
c	18	40.6	3.4	160482	72	AC055862 AC055862 Homo sapi
c	19	40.6	3.4	186558	79	AL162735 AL162735 Homo sapi
	20	39.4	3.3	7218	10	I66494 I66494 Sequence 14
	21	39.4	3.3	158802	63	AC015589 AC015589 Homo sapi
c	22	39.2	3.3	193260	69	AC025598 AC025598 Homo sapi
c	23	39	3.2	19020	6	CELR13H8 AF039717 Caenorhab
c	24	39	3.2	109140	83	AP001969 AP001969 Homo sapi
	25	39	3.2	155628	71	AC027810 AC027810 Homo sapi
	26	39	3.2	208515	68	AC024619 AC024619 Homo sapi
	27	38.6	3.2	141268	75	AC074365 AC074365 Homo sapi
	28	38.6	3.2	158475	83	CNS01DXB AL139194 Homo sapi
	29	38.4	3.2	1320	95	RNAPN1 AF039889 Rattus no
c	30	38.4	3.2	45408	77	AC087406 AC087406 Homo sapi
c	31	38.4	3.2	182382	76	AC079682 AC079682 Mus muscu
c	32	38	3.2	103270	13	ATF1N20 AL022140 Arabidops
c	33	38	3.2	196247	13	ATCHRIV56 AL161556 Arabidops
	34	37.8	3.1	132605	68	AC023450 AC023450 Homo sapi
c	35	37.8	3.1	210350	76	AC079536 AC079536 Mus muscu
c	36	37.6	3.1	158749	76	AC079614 AC079614 Homo sapi
c	37	37.6	3.1	202058	64	AC016913 AC016913 Homo sapi
c	38	37.4	3.1	148691	92	HS52D1 Z96811 Human DNA s
	39	37.4	3.1	162213	71	AC027674 AC027674 Homo sapi
c	40	37.4	3.1	165888	72	AC053506 AC053506 Homo sapi
c	41	37.4	3.1	189631	61	AC009920 AC009920 Homo sapi
	42	37.4	3.1	208342	66	AC020907 AC020907 Homo sapi
c	43	37.2	3.1	97375	90	AL353898 AL353898 Human DNA
	44	37.2	3.1	110000	60	AC008749_2 Continuation (3 of
c	45	37.2	3.1	154442	79	AL161915 AL161915 Homo sapi

ALIGNMENTS

```
RESULT   1
AX097623
LOCUS       AX097623     1203 bp    DNA             PAT       30-MAR-2001
DEFINITION  Sequence 3 from Patent WO0118051.
ACCESSION   AX097623
VERSION     AX097623.1  GI:13514264
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Oppmann,B., de waal Malefyt,R., Rennick,D.M., Kastelein,R.A.,
            Wiekowski,M.T., Lira,S.A. and Narula,S.K.
  TITLE     Mammalian interleukin-12 p40 and interleukin b30. Combinations
            thereof. Antibodies. Uses in pharmaceutical compositions
  JOURNAL   Patent: WO 0118051-A 3 15-MAR-2001;
            SCHERING CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="surmised Mus sp."
     CDS             113. .703
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35430.1"
                     /db_xref="GI:13514265"
                     /translation="MLDCRAVIMLWLLPWVTQGLAVPRSSSPDWAQCQQLSRNLCMLA
                     WNAHAPAGHMNLLREEEDEETKNNVPRIQCEDGCDPQGLKDNSQFCLQRIRQGLAFYK
                     HLLDSDIFKGEPALLPDSPMEQLHTSLLGLSQLLQPEDHPRETQQMPSLSSSQQWQRP
                     LLRSKILRSLQAFLAIAARVFAHGAATLTEPLVPTA"
     mat_peptide     176. .700
BASE COUNT      336 a    310 c    312 g    245 t
ORIGIN
```

Query Match 100.0%; Score 1203; DB 10; Length 1203;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60
Db       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60

Qy      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120
Db      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120

Qy     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180
Db     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180

Qy     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240
Db     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240

Qy     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300
Db     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300

Qy     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360
Db     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360

Qy     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
Db     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420

Qy     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480
Db     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480

Qy     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540
Db     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540

Qy     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600
Db     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600

Qy     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660
Db     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660

Qy     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720
Db     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720

Qy     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780
Db     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780

Qy     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840
Db     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840

Qy     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900
Db     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900

Qy     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
Db     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
```

```
Qy     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020

Qy    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080

Qy    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140

Qy    1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200

Qy    1201 AAA 1203
           |||
Db    1201 AAA 1203


RESULT   2
AF301619
LOCUS       AF301619     1358 bp    mRNA            ROD       04-DEC-2000
DEFINITION  Mus musculus interleukin 23 p19 subunit mRNA, complete cds.
ACCESSION   AF301619
VERSION     AF301619.1  GI:11528340
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1358)
  AUTHORS   Oppmann,B., Lesley,R., Blom,B., Timans,J.C., Xu,Y., Hunte,B.,
            Vega,F., Yu,N., Wang,J., Singh,K., Zonin,F., Vaisberg,E.,
            Churakova,T., Liu,M., Gorman,D., Wagner,J., Zurawski,S., Liu,Y.,
            Abrams,J.S., Moore,K.W., Rennick,D., de Waal-Malefyt,R., Hannum,C.,
            Bazan,J.F. and Kastelein,R.A.
  TITLE     Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with
            biological activities similar as well as distinct from IL-12
  JOURNAL   Immunity 13 (5), 715-725 (2000)
  MEDLINE   20567322
REFERENCE   2  (bases 1 to 1358)
  AUTHORS   Kastelein,R.A., Gorman,D., Timans,J.C., Oppmann,B. and Bazan,J.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) Molecular Biology, DNAX Research Institute,
            901 California Avenue, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .1358
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             113. .703
                     /note="four helical bundle cytokine; combines with the p40
                     subunit of IL-12 to form a novel, biologically active,
                     composite cytokine, IL-23"
                     /codon_start=1
                     /product="interleukin 23 p19 subunit"
                     /protein_id="AAG37231.1"
                     /db_xref="GI:11528341"
                     /translation="MLDCRAVIMLWLLPWVTQGLAVPRSSSPDWAQCQQLSRNLCMLA
                     WNAHAPAGHMNLLREEEDEETKNNVPRIQCEDGCDPQGLKDNSQFCLQRIRQGLAFYK
                     HLLDSDIFKGEPALLPDSPMEQLHTSLLGLSQLLQPEDHPRETQQMPSLSSSQQWQRP
                     LLRSKILRSLQAFLAIAARVFAHGAATLTEPLVPTA"
BASE COUNT      383 a    326 c    338 g    311 t
ORIGIN


  Query Match          100.0%;  Score 1203;  DB 94;  Length 1358;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60

Qy      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120

Qy     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180

Qy     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240

Qy     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300

Qy     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360

Qy     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420

Qy     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480

Qy     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540

Qy     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600

Qy     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660

Qy     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720

Qy     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780

Qy     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840

Qy     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900

Qy     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960

Qy     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020

Qy    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080

Qy    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
```

```
Qy   1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200

Qy   1201 AAA 1203
          |||
Db   1201 AAA 1203


RESULT   3
AF301620
LOCUS       AF301620     1025 bp    mRNA            PRI       04-DEC-2000
DEFINITION  Homo sapiens interleukin 23 p19 subunit mRNA, complete cds.
ACCESSION   AF301620
VERSION     AF301620.1  GI:11528342
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Oppmann,B., Lesley,R., Blom,B., Timans,J.C., Xu,Y., Hunte,B.,
            Vega,F., Yu,N., Wang,J., Singh,K., Zonin,F., Vaisberg,E.,
            Churakova,T., Liu,M., Gorman,D., Wagner,J., Zurawski,S., Liu,Y.,
            Abrams,J.S., Moore,K.W., Rennick,D., de Waal-Malefyt,R., Hannum,C.,
            Bazan,J.F. and Kastelein,R.A.
  TITLE     Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with
            biological activities similar as well as distinct from IL-12
  JOURNAL   Immunity 13 (5), 715-725 (2000)
  MEDLINE   20567322
REFERENCE   2  (bases 1 to 1025)
  AUTHORS   Kastelein,R.A., Gorman,D., Timans,J.C., Oppmann,B. and Bazan,J.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) Molecular Biology, DNAX Research Institute,
            901 California Avenue, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             151. .720
                     /note="combines with the p40 subunit of IL-12 to form a
                     novel, biologically active, composite cytokine, IL-23"
                     /codon_start=1
                     /product="interleukin 23 p19 subunit"
                     /protein_id="AAG37232.1"
                     /db_xref="GI:11528343"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVAQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
BASE COUNT      262 a    252 c    274 g    237 t
ORIGIN


  Query Match           36.7%;  Score 441.4;  DB 89;  Length 1025;
  Best Local Similarity  77.5%;  Pred. No. 4.5e-124;
  Matches  601;  Conservative    0;  Mismatches  161;  Indels   13;  Gaps    5;

Qy     18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
          | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db     55 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 114

Qy     77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
          | ||||||||||||||| ||||||| ||| || ||||||||||   ||||||| |||||
Db    115 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 174

Qy    137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
           |||| | |||| ||||||||    ||||||||   ||||||||| || | ||||| |||
Db    175 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 234

Qy    197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
          | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||

Db    235 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 294

Qy    257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
          ||| || |||||| ||| ||   |||||||||||| |  |||||||||||| |||| ||
Db    295 CCACTAGTGGGACACATGGAT---CTAAGAGAAGAGGGAGATGAAGAGACTACAAATGAT 351

Qy    317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
          || |||| |||||||||||| |||||| |||||||| |||||||||| ||||||||| |||
Db    352 GTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAG 411

Qy    377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
          |||||||||||||||||| ||| ||||||  |||||||| || | ||||| |  || ||
Db    412 TTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGAT 471

Qy    437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
          || |||| |||||||||| |||| ||||||||||||||  ||| ||| |||||  |||||
Db    472 ATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCC 531

Qy    497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
          ||||| || |||||||||||||||| |||| |||| |||| | ||||||| || ||||
Db    532 CTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATT 591

Qy    557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
          || ||||| ||| | || |||| ||||||| | |||||||| | ||| ||||||||
Db    592 CCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGC 651

Qy    617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
          |||||||||||||| |||| |||| |||||||||||||||| ||||||||||
Db    652 AGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAAC---- 707

Qy    677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
          ||||| || ||  | || ||||| |||||||    | ||| |  ||||||  |  | | |
Db    708 CCTGAGTCCCTAAAGGCAGCAGCTCAAGGATGGCACTCAGATCTCCATGGCCCAGCAAGG 767

Qy    732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
            ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db    768 CCAAGATAAATCTACCACCCCAGGCACCTGTGAGCCAACAGGTTAATTAGTCCAT 822


RESULT   4
AB030000
LOCUS       AB030000     1026 bp    mRNA            PRI       04-APR-2000
DEFINITION  Homo sapiens mRNA for SGRF, complete cds.
ACCESSION   AB030000
VERSION     AB030000.1  GI:7416072
KEYWORDS    SGRF.
SOURCE      Homo sapiens fetal spleen cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1026)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 1026)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .1026
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="fetal"
                     /tissue_type="spleen"
     gene            144. .713
                     /gene="SGRF"
     CDS             144. .713
                     /gene="SGRF"
                     /codon_start=1
```

```
                     /product="SGRF"
                     /protein_id="BAA93686.1"
                     /db_xref="GI:7416073"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVGQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
     sig_peptide     144. .200
                     /gene="SGRF"
     mat_peptide     201. .710
                     /gene="SGRF"
                     /product="SGRF"
     polyA_signal    996. .1001
BASE COUNT      268 a    249 c    272 g    237 t
ORIGIN
```

Query Match 36.6%; Score 439.8; DB 85; Length 1026;
Best Local Similarity 77.4%; Pred. No. 1.4e-123;
Matches 600; Conservative 0; Mismatches 162; Indels 13; Gaps 5;

```
Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      48 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 107

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||    ||||||| |||||
Db     108 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 167

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
            |||| | |||| ||||||||    |||||||   |||||||| || | ||||| |||
Db     168 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 227

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     228 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 287

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
            ||| || |||||| ||| ||   |||||||||| |  |||||||||||| |||| ||
Db     288 CCACTAGTGGGACACATGGAT---CTAAGAGAAGAGGGAGATGAAGAGACTACAAATGAT 344

Qy     317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
           || |||| |||||||||| |||||| |||||||| |||||||||  |||||||| |||
Db     345 GTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAG 404

Qy     377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
           ||||||||||||||||| ||| ||||||  |||||| || | ||||| |  || ||
Db     405 TTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGAT 464

Qy     437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
           || |||| |||||||||| |||| ||||||||||||||  |||  || |||||  |||||
Db     465 ATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCC 524

Qy     497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
           ||||| || |||||||||||||| ||||| |||| ||||| | ||||||| || |||||
Db     525 CTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATT 584

Qy     557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
           || ||||| ||| | || ||||  |||||||| | ||||||||| | ||| ||||||||
Db     585 CCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGC 644

Qy     617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
           |||||||||||||||| ||||  |||| |||||||||||||||| ||||||||||
Db     645 AGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAAC---- 700

Qy     677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
            ||||| || ||  | || ||||| |||||||   | ||| |  ||||||  |  | | |
Db     701 CCTGAGTCCCTAAAGGCAGCAGCTCAAGGATGGCACTCAGATCTCCATGGCCCAGCAAGG 760

Qy     732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
             ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db     761 CCAAGATAAATCTACCACCCCAGGCACCTGTGAGCCAACAGGTTAATTAGTCCAT 815
```

```
RESULT   5
AX048200
LOCUS       AX048200     1055 bp    DNA             PAT       15-DEC-2000
DEFINITION  Sequence 40 from Patent WO0070049.
ACCESSION   AX048200
VERSION     AX048200.1  GI:11876990
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1055)
  AUTHORS   Tang,Y.T., Yue,H., Lal,P., Burford,N., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Lu,D.A. and Patterson,C.
  TITLE     Extracellular signaling molecules
  JOURNAL   Patent: WO 0070049-A 40 23-NOV-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1055
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 2933038CB1"
BASE COUNT      284 a    254 c    278 g    239 t
ORIGIN
```

Query Match 36.6%; Score 439.8; DB 9; Length 1055;
Best Local Similarity 77.4%; Pred. No. 1.4e-123;
Matches 600; Conservative 0; Mismatches 162; Indels 13; Gaps 5;

```
Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      75 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 134

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||    ||||||| |||||
Db     135 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 194

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
            |||| | |||| ||||||||    |||||||   |||||||| || | ||||| |||
Db     195 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 254

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     255 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 314

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
            ||| || |||||| ||| ||   |||||||||| |  |||||||||||| |||| ||
Db     315 CCACTAGTGGGACACATGGAT---CTAAGAGAAGAGGGAGATGAAGAGACTACAAATGAT 371

Qy     317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
           || |||| |||||||||| |||||| |||||||| |||||||||  |||||||| |||
Db     372 GTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAG 431

Qy     377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
           ||||||||||||||||| ||| ||||||  |||||||  || | |||||  |  || ||
Db     432 TTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGAT 491

Qy     437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
           || |||| |||||||||| |||| ||||||||||||||  |||  || |||||  |||||
Db     492 ATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCC 551

Qy     497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
           ||||| || |||||||||||||| ||||| |||| ||||| | ||||||| || |||||
Db     552 CTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATT 611

Qy     557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
           || ||||| ||| | || ||||  |||||||| | ||||||||| | ||| ||||||||
Db     612 CCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGC 671

Qy     617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
           |||||||||||||||| ||||  |||| |||||||||||||||| ||||||||||
```

```
Db    672 AGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAAC---- 727

Qy    677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
          ||||| || ||  | || ||||| |||||||   | ||| |  ||||||  |  | | |
Db    728 CCTGAGTCCCTAAAGGCAGCAGCTCAAGGATGGCACTCAGATCTCCATGGCCCAGCAAGG 787

Qy    732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
          ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db    788 CCAAGATAAATCTACCACCCCAGGCACCTGTGAGCCAACAGGTTAATTAGTCCAT 842


RESULT   6
AX097621
LOCUS       AX097621      570 bp    DNA             PAT       30-MAR-2001
DEFINITION  Sequence 1 from Patent WO0118051.
ACCESSION   AX097621
VERSION     AX097621.1  GI:13514261
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Oppmann,B., de waal Malefyt,R., Rennick,D.M., Kastelein,R.A.,
            Wiekowski,M.T., Lira,S.A. and Narula,S.K.
  TITLE     Mammalian interleukin-12 p40 and interleukin b30. Combinations
            thereof. Antibodies. Uses in pharmaceutical compositions
  JOURNAL   Patent: WO 0118051-A 1 15-MAR-2001;
            SCHERING CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="surmised Homo sapiens"
     CDS             1. .570
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35429.1"
                     /db_xref="GI:13514262"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVAQLHASLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
     mat_peptide     64. .567
BASE COUNT      121 a    168 c    157 g    124 t
ORIGIN


  Query Match           32.5%;  Score 391;  DB 10;  Length 570;
  Best Local Similarity   81.8%;  Pred. No. 1.3e-108;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172
          |||||||   ||||||| |||||  |||| | |||| |||||||||    ||||||||
Db      1 ATGCTGGGGAGCAGAGCTGTAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGA 60

Qy    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232
          |||||||||| || | ||||| |||| |||| ||||||||||||||||| || | ||| |||
Db     61 GCTGTGCCTGGGGGCAGCAGCCCTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTC 120

Qy    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292
          ||||  || |||||||  |||||| ||| || |||||| ||| ||   ||||||||||
Db    121 TGCACACTGGCCTGGAGTGCACATCCACTAGTGGGACACATGGAT---CTAAGAGAAGAG 177

Qy    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352
          |  |||||||||||||| |||| |||| |||| |||||||||| |||||| ||||||||
Db    178 GGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGATGGCTGTGACCCC 237

Qy    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412
          |||||||||| ||||||||| ||||||||||||||||||||||| ||| ||||||  ||||
Db    238 CAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAGGGTCTGATTTTT 297

Qy    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||||

Db    298 TATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTGCTCCCTGATAGC 357

Qy    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532
          ||  ||| ||| |||||  |||||||||| || ||||||||||||| ||||| |||| |
Db    358 CCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTGCAGCCTGAGGGT 417

Qy    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592
          |||| | ||||||| || ||||| || ||||| ||| | || |||  ||||||||| | |
Db    418 CACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCATGGCAGCGTCTC 477

Qy    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652
          |||||||| | ||| |||||||| |||||||||||||| |||| |||| |||||||||
Db    478 CTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTAGCCGCCCGGGTC 537

Qy    653 TTTGCCCACGGAGCAGCAACTCTGACT 679
          |||||||| |||||||||| |||| |
Db    538 TTTGCCCATGGAGCAGCAACCCTGAGT 564


RESULT   7
AB030001
LOCUS       AB030001     4039 bp    DNA             PRI       04-APR-2000
DEFINITION  Homo sapiens gene for SGRF, complete cds.
ACCESSION   AB030001
VERSION     AB030001.1  GI:7416074
KEYWORDS    SGRF.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 4039)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q13"
     TATA_signal     38. .43
     exon            318. .622
                     /product="SGRF"
     gene            join(461. .622,842. .940,1107. .1253,1359. .1520)
                     /gene="SGRF"
     CDS             join(461. .622,842. .940,1107. .1253,1359. .1520)
                     /gene="SGRF"
                     /codon_start=1
                     /product="SGRF"
                     /protein_id="BAA93687.1"
                     /db_xref="GI:7416075"
                     /translation="MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLA
                     WSAHPLVGHMDLREEGDEETTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEK
                     LLGSDIFTGEPSLLPDSPVGQLHASLLGLSQLLQPEGHHWETQQ1PSLSPSQPWQRLL
                     LRFKILRSLQAFVAVAARVFAHGAATLSP"
     exon            842. .940
                     /gene="SGRF"
                     /product="SGRF"
     exon            1107. .1253
                     /gene="SGRF"
                     /product="SGRF"
     exon            1359. .1826
                     /product="SGRF"
     polyA_signal    1803. .1808
BASE COUNT     1098 a    940 c   1029 g    972 t
```

```
ORIGIN


  Query Match            12.6%;  Score 151.6;  DB 85;  Length 4039;
  Best Local Similarity   76.7%;  Pred. No. 5.1e-35;
  Matches  198;  Conservative    0;  Mismatches   59;  Indels    1;  Gaps    1;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db     365 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 424

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||||   ||||||| |||||
Db     425 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 484

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
           |||| | |||| ||||||||    ||||||||   ||||||||| || | ||||| |||
Db     485 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 544

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| ||||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     545 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 604

Qy     257 GCACCAGCGGGACATATG 274
           ||| || |||||| |||
Db     605 CCACTAGTGGGACACATG 622


RESULT   8
AC025574/c
LOCUS       AC025574   155929 bp    DNA             HTG       06-SEP-2000
DEFINITION  Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC025574
VERSION     AC025574.10  GI:9438393
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155929)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155929)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 25, 2000 this sequence version replaced gi:8571540.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: HALV
                Center clone name: RP11-348M3
            ---------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 6% of reads
                Chemistry: Dye-terminator Big Dye: 94% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 133622 bases at least Q40
                Consensus quality: 145733 bases at least Q30
                Consensus quality: 150334 bases at least Q20
                Estimated insert size: 150694; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    36972: contig of 36972 bp in length
            *    36973    37072: gap of unknown length
            *    37073    64901: contig of 27829 bp in length
            *    64902    65001: gap of unknown length
            *    65002    91014: contig of 26013 bp in length
            *    91015    91114: gap of unknown length
            *    91115   105122: contig of 14008 bp in length
            *   105123   105222: gap of unknown length
            *   105223   113963: contig of 8741 bp in length
            *   113964   114063: gap of unknown length
            *   114064   124894: contig of 10831 bp in length
            *   124895   124994: gap of unknown length
            *   124995   132612: contig of 7618 bp in length
            *   132613   132712: gap of unknown length
            *   132713   139659: contig of 6947 bp in length
            *   139660   139759: gap of unknown length
            *   139760   143764: contig of 4005 bp in length
            *   143765   143864: gap of unknown length
            *   143865   148062: contig of 4198 bp in length
            *   148063   148162: gap of unknown length
            *   148163   151000: contig of 2838 bp in length
            *   151001   151100: gap of unknown length
            *   151101   153374: contig of 2274 bp in length
            *   153375   153474: gap of unknown length
            *   153475   154810: contig of 1336 bp in length
            *   154811   154910: gap of unknown length
            *   154911   155929: contig of 1019 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-348M3"
BASE COUNT    40876 a  36387 c  36205 g  41119 t   1342 others
ORIGIN


  Query Match            12.5%;  Score 150.6;  DB 69;  Length 155929;
  Best Local Similarity   76.7%;  Pred. No. 1.6e-34;
  Matches  197;  Conservative    0;  Mismatches   59;  Indels    1;  Gaps    1;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db  64170 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 64111

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||||   ||||||| |||||
```

```
Db  64110 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 64051

Qy    137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
          |||| | |||| |||||||||    ||||||||  |||||||||| || | ||||| |||
Db  64050 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 63991

Qy    197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
          | |||| |||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db  63990 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 63931

Qy    257 GCACCAGCGGGACATAT 273
          ||| || |||||| ||
Db  63930 CCACTAGTGGGACACAT 63914


RESULT   9
AB030002
LOCUS       AB030002     2300 bp    DNA             MAM       04-APR-2000
DEFINITION  Sus scrofa gene for SGRF, complete cds.
ACCESSION   AB030002
VERSION     AB030002.1  GI:7416076
KEYWORDS    SGRF.
SOURCE      Sus scrofa DNA.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2300)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     SGRF; a novel member of the IL-6/G-CSF family
  JOURNAL   Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 2300)
  AUTHORS   Hirata,Y. and Kosuge,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) to the DDBJ/EMBL/GenBank databases. Yuichi
            Hirata, Chugai Research Institute for Molecular Medicine, Gene
            Search Program; 153-2 Nagai, Niihari-mura, Ibaraki 300-4101, Japan
            (E-mail:hiratayu@cimmed.com, Tel:81-298-30-6211(ex.288),
            Fax:81-298-30-6270)
FEATURES             Location/Qualifiers
     source          1. .2300
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     TATA_signal     227. .232
     gene            join(433. .603,820. .921,1590. .1736,1837. .1998)
                     /gene="SGRF"
     CDS             join(433. .603,820. .921,1590. .1736,1837. .1998)
                     /gene="SGRF"
                     /codon_start=1
                     /product="SGRF"
                     /protein_id="BAA93688.1"
                     /db_xref="GI:7416077"
                     /translation="MLGSRAVMLMLLLLLLPWTSQGRAVPEGSSPAWAQGQQLSQQLC
                     TLAWTAHLPMGHVDLPREEGDDETTSEVPHIQCGDGCDPQGLRDNSQSCLQRIHQGLV
                     FYEKLLGSDIFTGEPSLHPDGSVGQLHASLLGLRQLLQPEGHHWETEQTPSPSPSQPW
                     QRLLLRLKILRSLQAFVAVAARVFAHGAATLSQ"
     polyA_signal    2265. .2270
BASE COUNT      597 a    563 c    589 g    551 t
ORIGIN


  Query Match           10.2%;  Score 123.2;  DB 7;  Length 2300;
  Best Local Similarity   72.5%;  Pred. No. 2.5e-26;
  Matches  190;  Conservative    0;  Mismatches   63;  Indels    9;  Gaps    2;

Qy    517 CCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCA 576
          | | |||||  |||| |||||  | |||||||  | ||||  ||| ||||  ||| | || ||
Db   1830 CTTTCAGCCCGAGGGTCACCACTGGGAGACTGAGCAGACGCCAAGCCCCAGTCCCAGCCA 1889

Qy    577 GCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGC 636
          ||  ||||| |||| |||||||||   ||||||||||| ||||||||||||||| ||||
Db   1890 GCCCTGGCAACGCCTCCTTCTCCGCCTCAAGATCCTTCGCAGCCTCCAGGCCTTTGTGGC 1949

Qy    637 CATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAAC 696
          |||||||||||||||| ||||| ||||||||||||    |||||||  ||  |||| |
Db   1950 TGTAGCTGCCCGGGTCTTCGCCCATGGAGCAGCAACT----CTGAGCCAGTAAAGCCAGC 2005

Qy    697 AGCTTAAGGATGCC-----CAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCC 751
          |||||||||||| |     |  | |     |  | ||| ||| |  || |||| ||||
Db   2006 AGCTTAAGGATGACACCAAGACCTCCATTGCTCAGCAATGCTAAAATGAAGCTATGAGCC 2065

Qy    752 CAGACATCTACCAGTTAATTAA 773
          | |  |  | ||||  |||||
Db   2066 AACAAGTTCATCAGTCCATTAA 2087


RESULT  10
AB004061/c
LOCUS       AB004061     4649 bp    mRNA            MAM       05-FEB-1999
DEFINITION  domestic pig mRNA for STAT2, complete cds.
ACCESSION   AB004061
VERSION     AB004061.1  GI:2189979
KEYWORDS    STAT2.
SOURCE      Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 4649)
  AUTHORS   Ito,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1997) to the DDBJ/EMBL/GenBank databases.
            Yoshiyasu Ito, National Institute of Animal Industry, Animal Genome
            Research Team; 2 Ikenodai, Kukizaki-machi, Inashiki-gun, Ibaragi
            305-0901, Japan (E-mail:yoshiito@niai.affrc.go.jp,
            Tel:81-298-38-8627)
REFERENCE   2  (sites)
  AUTHORS   Ito,Y., MIKAWA,S., Kobayashi,E., Wada,Y. and MINEZAWA,M.
  TITLE     Pig STAT2 cDNA
  JOURNAL   Unpublished (1997)
FEATURES             Location/Qualifiers
     source          1. .4649
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /cell_type="thoracic aortic cell"
                     /dev_stage="adult"
                     /tissue_type="muscle"
     CDS             82. .2676
                     /codon_start=1
                     /product="STAT2"
                     /protein_id="BAA20332.1"
                     /db_xref="GI:2189980"
                     /translation="MAQWEMLQNLDSPFQDQLHQLYSESLLPVDVRQYLAVWIEDQNW
                     QEAALGNDDSKANMLFFHFLDQLNYDCGRCGQDPECLLLQHNLRKFYRDIQAIPQGPT
                     RLAEMIFNLLLEEKRILIQAQRAQLEQQEPALEAPGENQQHEIESRILELRAMMEKLV
                     KSISQLKDQQDIFCFRYKIKASAKTHSLDPHRTRQEQVLQETLNELDKRRKEVLDASK
                     ALLGRLTTLIELLLPKLEEWKVQQQKACIGAPMDGELEQLEKWFTAEAKLLFHLRQLL
                     KELKGLSSVVKYDEDLLFKGVDLLKAQVTELLQRLLHRAFIVETQPCMPQTPHRPLIL
                     KTGSKFTVRTRLLVRLQEGNESLTAEVSIDRNPPKSQGFRKFNILTSNRKTLTPEKGQ
                     SQGLIWDFGYLTLMEQRSGGAGKGNNKGPLGVTEELHIISFTVKYTYQGLKQELTTDT
                     LPVVIISNMNQLSIAWASILWFNLLSSDPQNQQFFSSPPKAPWNLLGPALSWQFSSHV
                     GQGLNSDQLGMLRDKLFGQNSSTEGLSLSWVDFIKRESPPGKLPFWTWLDKILDLVHD
                     HLKDLWKDGHIMGFVSRSEERRLLKKTISGTFLLRFSETLEGGITCSWVEHQDDDKVL
                     IYSLQPFTKEVLQSLPLTKIISQYQLLTEENIPENPLRFLYPRIPRDEAFGCYNQEKA
                     NLQERKKYLKHKLIVVSNRQVDELQQPPELKLEPDLESLELDLGLAPGPEPGVGLDLE
                     PLLEAGLDLGPELEPMLQSTLEPVLEPILDEVLQGVPEPNLGPELKLEPILEPVSQPV
                     PEPDLPHDLRHLNTDEMQIFRNSMRIEEIMPNGDPLLASQNTNVDEACVFHRSHFYTD
                     GPLIPSDY"
     polyA_site      3622
                     /note="Alternative Poly adenilation site"
     polyA_site      4649
                     /note="21 a nucleotides"
BASE COUNT     1234 a   1157 c   1194 g   1064 t
ORIGIN


  Query Match           10.2%;  Score 123.2;  DB 7;  Length 4649;
  Best Local Similarity   72.5%;  Pred. No. 2.7e-26;
```

```
  Matches  190;  Conservative    0;  Mismatches   63;  Indels    9;  Gaps    2;

Qy     517 CCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCA 576
           | | ||||| |||| ||||| | ||||||| | |||| ||| |||| ||| | || ||
Db    3866 CTTTCAGCCCGAGGGTCACCACTGGGAGACTGAGCAGACGCCAAGCCCCAGTCCCAGCCA 3807

Qy     577 GCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGC 636
           ||  ||||| |||| ||||||||||   |||||||||||| |||||||||||||| ||||
Db    3806 GCCCTGGCAACGCCTCCTTCTCCGCCTCAAGATCCTTCGCAGCCTCCAGGCCTTTGTGGC 3747

Qy     637 CATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAAC 696
             |||||||||||||||| ||||| ||||||||||||    |||||||  ||  |||| |
Db    3746 TGTAGCTGCCCGGGTCTTCGCCCATGGAGCAGCAACT----CTGAGCCAGTAAAGCCAGC 3691

Qy     697 AGCTTAAGGATGCC-----CAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCC 751
           |||||||||||| |      |  | |      |  | ||| ||| |  || |||| ||||
Db    3690 AGCTTAAGGATGACACCAAGACCTCCATTGCTCAGCAATGCTAAAATGAAGCTATGAGCC 3631

Qy     752 CAGACATCTACCAGTTAATTAA 773
           | |  |  | ||||  |||||
Db    3630 AACAAGTTCATCAGTCCATTAA 3609


RESULT  11
I66494/c
LOCUS       I66494       7218 bp    DNA             PAT       28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .7218
                     /organism="unknown"
BASE COUNT     1944 a   1491 c   1486 g   1929 t    368 others
ORIGIN


  Query Match             5.2%;  Score 62;  DB 10;  Length 7218;
  Best Local Similarity   2.5%;  Pred. No. 1.8e-07;
  Matches    8;  Conservative  200;  Mismatches  110;  Indels    0;  Gaps    0;

Qy     816 GAAGGGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCCAGAGCAGGCAGCTGGC 875
           :::::: ::::: :  :  :  :::::::::::::: :::    :::: ::: ::  ::
Db    1370 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1311

Qy     876 TAGAGAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCA 935
            ::::::::::  :::::::::::: :::    :::     ::    ::::: ::: :
Db    1310 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1251

Qy     936 AGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCA 995
           :: ::  : : ::   :::::::::::::: :: :: :  ::::::  : :::::::   :
Db    1250 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1191

Qy     996 GAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGG 1055
           ::: :  :::::::: :::   :::    ::  :  :   : :::::::: :::::::::
Db    1190 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1131

Qy    1056 GGAACATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACA 1115
           :::: :  : :       ::: ::   ::::::: : : ::: : : :: :    ::: :
Db    1130 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1071

Qy    1116 GCAGCTCTGTACCTGCCT 1133
           : :: ||   | || |||
Db    1070 RRRRATCGCAAGCTCCCT 1053


RESULT  12
NCB24M22/c
LOCUS       NCB24M22    90558 bp    DNA             PLN       24-JUL-2000
DEFINITION  Neurospora crassa DNA linkage group II BAC contig B24M22.
ACCESSION   AL390354
VERSION     AL390354.1  GI:9453800
KEYWORDS    .
SOURCE      Neurospora crassa.
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariales; Sordariaceae; Neurospora.
REFERENCE   1  (bases 1 to 90558)
  AUTHORS   Schulte,U., Aign,V., Hoheisel,J., Brandt,P., Fartmann,B.,
            Holland,R., Nyakatura,G., Mewes,H.W. and Mannhaupt,G.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 90558)
  AUTHORS   German Neurospora genome,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            mannhaup@mips.biochem.mpg.de Project Coordinator: Ulrich Schulte,
            Institute of Biochemistry, Heinrich-Heine-University, D-40225
            Duesseldorf, E-mail: ulrich-schulte@uni-duesseldorf.de
COMMENT     this contig is an assembly of BAC B24M22 from 1 to: 48331 rev,
            strain OR74A, and cosmid 6a13 from 1 to: 17504 rev, cosmid library
            pMOcosX, strain 74-OR-23-1VA; BAC clones are available at the
            Fungal Genetic Stock Center, http://www.fgsc.net; cosmid clone 6A13
            is available from J. Hoheisel at DKFZ, e-mail:
            J.Hoheisel@dkfz-heidelberg.de Sequencing was performed by MWG
            Biotech AG, Ebersberg, Germany, http://www.mwgdna.com
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of linkage groups II
            and V can be viewed at:
            http://www.mips.biochem.mpg.de/proj/neurospora.
FEATURES             Location/Qualifiers
     source          1. .90558
                     /organism="Neurospora crassa"
                     /db_xref="taxon:5141"
                     /chromosome="6"
     misc_feature    1. .28402
                     /note="overlap to BAC B11O9"
     repeat_region   84. .115
                     /note="32 bp tcggtcggtcggtcggtcgg tandem repeat"
     gene            4012. .6722
                     /gene="B24M22.10"
     CDS             join(4012. .4071,4222. .5321,5382. .5497,5560. .5791,
                     5852. .6722)
                     /gene="B24M22.10"
                     /note="Contains Fungal Zn(2)-Cys(6) binuclear cluster
                     domain [ACLVCRKKRTKCDGQMPCRRCRSRGEEC]"
                     /codon_start=1
                     /product="probable C6 zinc cluster transcription factor
                     fluffy"
                     /protein_id="CAB99369.1"
                     /db_xref="GI:9453801"
                     /translation="MPRQHLTPNACLVCRKKRTKCDGQMPCRRCRSRGEECAYEDKKW
                     RTKDHLRSEIERLRNEQRQGHAVIRALINDEQDWESFLSRIRGDESPEAIADWIRSIR
                     NLFEPLQAASSQSMGGLGAPPTLLSPSQATASESSQLHRAASFAGIGSYNFGQGRVPF
                     DQSTPRSSFSSDLSPTTPFSFREQADFIHAPQPMYPSSRRFSSSSLPSLPLRHSSQPL
                     VPGIFNEPLPHTWTSITSDTQLVQRLLSRFFSAPCSLLCFIPQSSFMKAFREGDSRYC
                     SEALVNAILGKACKSYGTASNIVSRMAFGDAFIGEAKRLLATEPNHTNLPSTQALAVL
                     ALAEISEGKDDEAWDLAWASVRAAITREQSFHVDQEFATARAVSYCGGFTLIHMLRLL
                     TGRLDLNTSPFFMRLYQGSEETPEDEPQNRIERGFALHMQFLAELEHCPPLPRFVFEI
                     TTAVHTFASYNFSNAATAEELEDAYGKCLDAYKRFEETFCLDMDTTPDLLFAQIWYHY
                     CLLALLRPFVKSTASLRDSAMTTPRLRNDANPSDICQRSSEAIIFLTSTYQTRFSLGN
                     PPELLPHMLFAAVLYQVTLTPDPEHLSTIANDIKPELSESPVMMPSQAAFGAHGNSNL
                     VPPPPMPFNNHGSYFPQPLSPVLKLEVRQAAPRRESSISLSSTFDSCGNRRPSDSFTS
                     STLTSHDASERESSTSDTQSDFLPFFTSEPADLVTIGSLQLASMQHHGAVEATRLLRS
                     LSTVKDLVGSTLDLETLAEALPFPMGDLNTAVLYTGLGLQRAPVEPMQVTGP"
     exon            4012. .4071
                     /gene="B24M22.10"
                     /number=1
     intron          4072. .4221
```

```
                    /gene="B24M22.10"
                    /number=1
     misc_feature   4204. .4208
                    /gene="B24M22.10"
                    /note="lariat consensus CTAAC"
     exon           4222. .5321
                    /gene="B24M22.10"
                    /number=2
     intron         5322. .5381
                    /gene="B24M22.10"
                    /number=2
     misc_feature   5362. .5366
                    /gene="B24M22.10"
                    /note="lariat consensus CTGAC"
     exon           5382. .5497
                    /gene="B24M22.10"
                    /number=3
     intron         5498. .5559
                    /gene="B24M22.10"
                    /number=3
     misc_feature   5506. .5510
                    /gene="B24M22.10"
                    /note="lariat consensus CTAAT"
     exon           5560. .5791
                    /gene="B24M22.10"
                    /number=4
     intron         5792. .5851
                    /gene="B24M22.10"
                    /number=4
     misc_feature   5835. .5839
                    /gene="B24M22.10"
                    /note="lariat consensus CTGAC"
     exon           5852. .6722
                    /gene="B24M22.10"
                    /number=5
     gene           10806. .11531
                    /gene="B24M22.20"
     CDS            10806. .11531
                    /gene="B24M22.20"
                    /note="weak similarity to proline-rich protein A316R -
                    Chlorella virus PBCV-1, PIR2:T17815"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAB99370.1"
                    /db_xref="GI:9453802"
                    /translation="MPSRQELAAPSSAPTPAPAPAPAPSNKKKPKKITGPPPHSEPFK
                    RPYHVSKDRPMSWYRGDANRIFLWSPFSKAFPLLLLDKLPKLQPRPISDDDEPIAVNG
                    FDATKKHFYAAQAKWERAVEIMEEEGRRMVYYGRTMHEEAQACHKWVGQVGKMTWEGH
                    HGRMKNYDAFEEALDKVCKMNKKAVETVEEVQESTKKWPVDLYARGGLGRKKRKTRPT
                    TMKRTESKRWLLCSFLGFKVTTS"
     exon           10806. .11531
                    /gene="B24M22.20"
                    /number=1
     exon           12310. .12636
                    /gene="B24M22.30"
                    /number=1
     gene           12310. .12636
                    /gene="B24M22.30"
     CDS            12310. .12636
                    /gene="B24M22.30"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAB99371.1"
                    /db_xref="GI:9453803"
                    /translation="MQLLHFLVFIFTTTVAASNSSCFDDGNFQQSCDSDSFSFTPHMG
                    ISAVCQRGDAESNLTTTLGLSKCIAYNIKESKLEWKIPPSCTNSTTVDEKKMNSTQGN
                    LSVLAS"
     gene           13650. .13997
                    /gene="B24M22.40"
     exon           13650. .13997
                    /gene="B24M22.40"
                    /number=1
     CDS            13650. .13997
                    /gene="B24M22.40"
                    /note="weak similarity to homoserine kinase (EC 2.7.1.39)
                    - Pseudomonas aeruginosa, PIR1:S27981"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAB99372.1"
                    /db_xref="GI:9453804"
                    /translation="MHFPLLPPTLAHILTTLATLIPTISATTIANTDPSQKPMIDTFR
                    KSCVNWQVDSKACALSGICKLTDDSHQASHPKLHGWRETNLDLNQCIWFDAKESKLKW
                    FEEKDGKVEDLSG"
     CDS            join(15172. .15664,15730. .17543)
                    /gene="B24M22.50"
                    /note="similarity to tol protein, Neurospora crassa,
                    PIR:T17430"
                    /codon_start=1
                    /product="related to tol protein"
                    /protein_id="CAB99373.1"
                    /db_xref="GI:9453805"
                    /translation="MTLDGGVIPHVLCYRCSRIVHSDWVQERLARLEPWTGTSTWALD
                    GDAGQEEKLRFFHSESLEKLEASAAAGCHLCTLISKKNQYFEWFKDEEHAGVVQVVAC
                    AGHSSVWLETSRIQELEIEEDTPHDLSGLDDWLEERYNQSLDFSGAWIMDDFVTVQLR
                    KVQSDTYSPVGSESNLPLSASSDAPQVLESIQCWLKDCVDNHPECNRYSSTQKDPTAT
                    WLPTRLIDVGEDPGKLAPRLVISADEGLAQKKAHYLTLSHSWSISTKPHNLKLETENM
                    TQLQDRIPLDDRLSKTFRDAMKITQQLGCRYIWIDSLCIIQSGIKAGDDWKREGLTMS
                    DVYGHSRCNITALGTGGNDGCYTPRNPLQTSPCHIKVCDDKPAIYAVNSLLENTFTEN
                    VEQSPLFKRGWVFQESVLSSRAVYFGAPQLFWECRRQAFAESWPFKTGTPGVQQRLTE
                    PTLRATGDTPFSSKTVFEALCNAEEQQTDGSHEQNDREKHLIGHTFPSFLCWEKIVKT
                    FIGTSLTFTSDRLFAMAGIARAIEHFRSFTYVTGSWRELYPLDLLWRFTRELGERKNV
                    VSGSKAPSWSWAAKELDMQESLKNIARSGKNTQRPAHWDGNKRFLVESVTSVAYLSQL
                    IKFPEPEFESKYRPDKEGEITIDLKGRVCKGLVKTELDIQGFPQDRLYAAGKDKAPVI
                    VHMDDDTPADGLEVVLVLLMEWEGLVGNYWKRQFQAGLVLVPAKDEGDKGGSTIYRRV
                    GAFRTGELSECPVNAFDGSEMEEVLGIC"
     exon           15172. .15664
                    /gene="B24M22.50"
                    /number=1
     gene           15172. .17543
                    /gene="B24M22.50"
     intron         15665. .15729
                    /gene="B24M22.50"
                    /number=1
     misc_feature   15711. .15715
                    /gene="B24M22.50"
                    /note="lariat consensus CTGAC"
     exon           15730. .17543
                    /gene="B24M22.50"
                    /number=2
     gene           19570. .20185
                    /gene="B24M22.60"
     CDS            join(19570. .19911,19982. .20185)
                    /gene="B24M22.60"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAB99374.1"
                    /db_xref="GI:9453806"
                    /translation="MVNFTRAAAALALAATATAAPTNGTTSLTWDVKDFEFKAEYQFT
                    TPAHQNSYGHVNFSLTSEKTGAVYSCGAQSSQIPDFFYDFNDFYPCSAPYGAVEGTES

  Query Match             3.6%;  Score 43.8;  DB 14;  Length 90558;
  Best Local Similarity   50.7%;  Pred. No. 0.092;
  Matches  105;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy     700 TTAAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCCAGACATC 759
           | |||||||||||| | || |||| | |   |    | || || || | | | || ||
Db   84330 TGAAGGATGCCCAAGATCTCATGTCCAAGCTACACAAACAAAACTTTGACCTCAAACTGG 84271

Qy     760 TACCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAG 819
            ||   || ||   | |   | |     ||| |   |     |||    ||  |||
Db   84270 AACTCTTTCATCGGCGCGAGAAGCAGACGGCTCTGGAGGAGCTAATTGAGAAGCTGGAAC 84211

Qy     820 GGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGA 879
           | |   | |||| |     || ||  | |   | || |  || |     |  | || |
Db   84210 TGGAGAAAGACCAGTTGAGGAGGGAGCAAAGGGAGACACAGGACCTCAATGACGTCTTGT 84151
```

```
Qy    880 GAAAGGAGCTGGAGAAGAAGAATAAAG 906
          | ||||||||||||||| | || |
Db  84150 TGGAAGAGCTGGAGAAGAAGGACAAGG 84124


RESULT  13
AC020877/c
LOCUS       AC020877   258622 bp    DNA             HTG       29-MAR-2000
DEFINITION  Mus musculus clone RP23-344L20, WORKING DRAFT SEQUENCE, 97
            unordered pieces.
ACCESSION   AC020877
VERSION     AC020877.2  GI:7340303
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 258622)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 258622)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 29, 2000 this sequence version replaced gi:6686431.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------Summary Statistics
            Consensus quality: 140189 bases at least Q40
            Consensus quality: 191513 bases at least Q30
            Consensus quality: 208421 bases at least Q20
            Estimated insert size: 258622; sum-of-contigs estimation
            Estimated insert size: 216000; pulse field gel estimation
            Quality coverage:  4.19x in Q20 bases; pulse field gel estimation
            Quality coverage:  3.50x in Q20 bases; sum-of-contigs estimation
            -------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 97 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1069: contig of 1069 bp in length
            *                     gap of unknown length
            *     1070     2085: contig of 1016 bp in length
            *                     gap of unknown length
            *     2086     3153: contig of 1068 bp in length
            *                     gap of unknown length
            *     3154     4722: contig of 1569 bp in length
            *                     gap of unknown length
            *     4723     5728: contig of 1006 bp in length
            *                     gap of unknown length
            *     5729     6764: contig of 1036 bp in length
            *                     gap of unknown length
            *     6765     7830: contig of 1066 bp in length
            *                     gap of unknown length
            *     7831     8859: contig of 1029 bp in length
            *                     gap of unknown length
            *     8860     9865: contig of 1006 bp in length
            *                     gap of unknown length
            *     9866    10869: contig of 1004 bp in length
            *                     gap of unknown length
            *    10870    11962: contig of 1093 bp in length
            *                     gap of unknown length
            *    11963    13013: contig of 1051 bp in length
            *                     gap of unknown length
            *    13014    14338: contig of 1325 bp in length
            *                     gap of unknown length
            *    14339    15557: contig of 1219 bp in length
            *                     gap of unknown length
            *    15558    16667: contig of 1110 bp in length
            *                     gap of unknown length
            *    16668    17748: contig of 1081 bp in length
            *                     gap of unknown length
            *    17749    19089: contig of 1341 bp in length
            *                     gap of unknown length
            *    19090    20675: contig of 1586 bp in length
            *                     gap of unknown length
            *    20676    22436: contig of 1761 bp in length
            *                     gap of unknown length
            *    22437    23683: contig of 1247 bp in length
            *                     gap of unknown length
            *    23684    25215: contig of 1532 bp in length
            *                     gap of unknown length
            *    25216    26431: contig of 1216 bp in length
            *                     gap of unknown length
            *    26432    27491: contig of 1060 bp in length
            *                     gap of unknown length
            *    27492    28737: contig of 1246 bp in length
            *                     gap of unknown length
            *    28738    29742: contig of 1005 bp in length
            *                     gap of unknown length
            *    29743    30947: contig of 1205 bp in length
            *                     gap of unknown length
            *    30948    31992: contig of 1045 bp in length
            *                     gap of unknown length
            *    31993    33108: contig of 1116 bp in length
            *                     gap of unknown length
            *    33109    34435: contig of 1327 bp in length
            *                     gap of unknown length
            *    34436    35651: contig of 1216 bp in length
            *                     gap of unknown length
            *    35652    37355: contig of 1704 bp in length
            *                     gap of unknown length
            *    37356    38598: contig of 1243 bp in length
            *                     gap of unknown length
            *    38599    39644: contig of 1046 bp in length
            *                     gap of unknown length
            *    39645    41275: contig of 1631 bp in length
            *                     gap of unknown length
            *    41276    42381: contig of 1106 bp in length
            *                     gap of unknown length
            *    42382    43462: contig of 1081 bp in length
            *                     gap of unknown length
            *    43463    44822: contig of 1360 bp in length
            *                     gap of unknown length
            *    44823    45961: contig of 1139 bp in length
            *                     gap of unknown length
            *    45962    47534: contig of 1573 bp in length
            *                     gap of unknown length
            *    47535    49241: contig of 1707 bp in length
            *                     gap of unknown length
            *    49242    50343: contig of 1102 bp in length
            *                     gap of unknown length
            *    50344    51833: contig of 1490 bp in length
            *                     gap of unknown length
            *    51834    52901: contig of 1068 bp in length
            *                     gap of unknown length
            *    52902    54595: contig of 1694 bp in length
            *                     gap of unknown length
            *    54596    56094: contig of 1499 bp in length
            *                     gap of unknown length
            *    56095    57171: contig of 1077 bp in length
            *                     gap of unknown length
            *    57172    58192: contig of 1021 bp in length
            *                     gap of unknown length
            *    58193    59434: contig of 1242 bp in length
            *                     gap of unknown length
            *    59435    60648: contig of 1214 bp in length
            *                     gap of unknown length
```

```
          *    60649    62130: contig of 1482 bp in length
          *                    gap of unknown length
          *    62131    63276: contig of 1146 bp in length
          *                    gap of unknown length
          *    63277    65248: contig of 1972 bp in length
          *                    gap of unknown length
          *    65249    67624: contig of 2376 bp in length
          *                    gap of unknown length
          *    67625    69132: contig of 1508 bp in length
          *                    gap of unknown length
          *    69133    71363: contig of 2231 bp in length
          *                    gap of unknown length
          *    71364    72844: contig of 1481 bp in length
          *                    gap of unknown length
          *    72845    74275: contig of 1431 bp in length
          *                    gap of unknown length
          *    74276    75619: contig of 1344 bp in length
          *                    gap of unknown length
          *    75620    76851: contig of 1232 bp in length
          *                    gap of unknown length
          *    76852    78211: contig of 1360 bp in length
          *                    gap of unknown length
          *    78212    80704: contig of 2493 bp in length
          *                    gap of unknown length
          *    80705    81729: contig of 1025 bp in length
          *                    gap of unknown length
          *    81730    84042: contig of 2313 bp in length
          *                    gap of unknown length
          *    84043    85320: contig of 1278 bp in length
          *                    gap of unknown length
          *    85321    87117: contig of 1797 bp in length
          *                    gap of unknown length
          *    87118    88340: contig of 1223 bp in length
          *                    gap of unknown length
          *    88341    89484: contig of 1144 bp in length
          *                    gap of unknown length
          *    89485    90573: contig of 1089 bp in length
          *                    gap of unknown length
          *    90574    93019: contig of 2446 bp in length
          *                    gap of unknown length
          *    93020    94631: contig of 1612 bp in length
          *                    gap of unknown length
          *    94632    96750: contig of 2119 bp in length
          *                    gap of unknown length
          *    96751    98951: contig of 2201 bp in length
          *                    gap of unknown length
          *    98952   100658: contig of 1707 bp in length
          *                    gap of unknown length
          *   100659   102977: contig of 2319 bp in length
          *                    gap of unknown length
          *   102978   106226: contig of 3249 bp in length
          *                    gap of unknown length
          *   106227   107977: contig of 1751 bp in length
          *                    gap of unknown length
          *   107978   110243: contig of 2266 bp in length
          *                    gap of unknown length
          *   110244   113543: contig of 3300 bp in length
          *                    gap of unknown length
          *   113544   118028: contig of 4485 bp in length
          *                    gap of unknown length
          *   118029   120548: contig of 2520 bp in length

  Query Match             3.6%;  Score 43.8;  DB 66;  Length 258622;
  Best Local Similarity   64.6%;  Pred. No. 0.1;
  Matches   82;  Conservative    0;  Mismatches   42;  Indels    3;  Gaps    1;

Qy     954 GGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGA 1013
           | ||   || |||| |   |  |||||  |||| ||  |||| ||  ||||| ||||
Db  140521 GTGGTGCGGAGCGGAGCTTTAAAGAAAGAGTGAAAAGGCCCAAAGTCTCAGAGAAAGCAG 140462

Qy    1014 GCCCAGGCTTTGGCCATTATCTGTAAGAAAAACA---AGAAAAGGGGAACATTATACTTT 1070
           ||  ||||| ||||||  ||||| |||| | |     |||||| ||||  | || | |
Db  140461 GCAGAGGCTCTGGCCAGCATCTGAAAGATAGAGGTGGAGAAAATGGGAAATTCATGCCTC 140402

Qy    1071 CCTGGGT 1077
           ||| ||
Db  140401 TCTGAGT 140395
```

RESULT 14
AC087468/c
LOCUS AC087468 62336 bp DNA HTG 05-JAN-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-255M6 map 15, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC087468
VERSION AC087468.1 GI:12039481
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62336)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-255M6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 62336)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
Center project name: L11910
Center clone name: 255_M_6
--------------.
* NOTE: This record contains 76 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 700: contig of 700 bp in length
* 701 800: gap of 100 bp
* 801 1526: contig of 726 bp in length

```
*     1527 1626: gap of      100 bp
*     1627     2364: contig of 738 bp in length
*     2365 2464: gap of      100 bp
*     2465     3198: contig of 734 bp in length
*     3199 3298: gap of      100 bp
*     3299     3959: contig of 661 bp in length
*     3960 4059: gap of      100 bp
*     4060     4778: contig of 719 bp in length
*     4779 4878: gap of      100 bp
*     4879     5619: contig of 741 bp in length
*     5620 5719: gap of      100 bp
*     5720     6435: contig of 716 bp in length
*     6436 6535: gap of      100 bp
*     6536     7261: contig of 726 bp in length
*     7262 7361: gap of      100 bp
*     7362     8081: contig of 720 bp in length
*     8082 8181: gap of      100 bp
*     8182     8914: contig of 733 bp in length
*     8915 9014: gap of      100 bp
*     9015     9744: contig of 730 bp in length
*     9745 9844: gap of      100 bp
*     9845    10570: contig of 726 bp in length
*    10571 10670: gap of      100 bp
*    10671    11396: contig of 726 bp in length
*    11397 11496: gap of      100 bp
*    11497    12222: contig of 726 bp in length
*    12223 12322: gap of      100 bp
*    12323    13050: contig of 728 bp in length
*    13051 13150: gap of      100 bp
*    13151    13865: contig of 715 bp in length
*    13866 13965: gap of      100 bp
*    13966    14675: contig of 710 bp in length
*    14676 14775: gap of      100 bp
*    14776    15517: contig of 742 bp in length
*    15518 15617: gap of      100 bp
*    15618    16330: contig of 713 bp in length
*    16331 16430: gap of      100 bp
*    16431    17142: contig of 712 bp in length
*    17143 17242: gap of      100 bp
*    17243    17967: contig of 725 bp in length
*    17968 18067: gap of      100 bp
*    18068    18783: contig of 716 bp in length
*    18784 18883: gap of      100 bp
*    18884    19604: contig of 721 bp in length
*    19605 19704: gap of      100 bp
*    19705    20428: contig of 724 bp in length
*    20429 20528: gap of      100 bp
*    20529    21215: contig of 687 bp in length
*    21216 21315: gap of      100 bp
*    21316    22044: contig of 729 bp in length
*    22045 22144: gap of      100 bp
*    22145    22875: contig of 731 bp in length
*    22876 22975: gap of      100 bp
*    22976    23696: contig of 721 bp in length
*    23697 23796: gap of      100 bp
*    23797    24523: contig of 727 bp in length
*    24524 24623: gap of      100 bp
*    24624    25337: contig of 714 bp in length
*    25338 25437: gap of      100 bp
*    25438    26169: contig of 732 bp in length
*    26170 26269: gap of      100 bp
*    26270    26978: contig of 709 bp in length
*    26979 27078: gap of      100 bp
*    27079    27772: contig of 694 bp in length
*    27773 27872: gap of      100 bp
*    27873    28599: contig of 727 bp in length
*    28600 28699: gap of      100 bp
*    28700    29405: contig of 706 bp in length
*    29406 29505: gap of      100 bp
*    29506    30231: contig of 726 bp in length
*    30232 30331: gap of      100 bp
*    30332    31052: contig of 721 bp in length
*    31053 31152: gap of      100 bp
*    31153    31884: contig of 732 bp in length
*    31885 31984: gap of      100 bp
*    31985    32698: contig of 714 bp in length
*    32699 32798: gap of      100 bp
*    32799    33519: contig of 721 bp in length
*    33520 33619: gap of      100 bp
*    33620    34367: contig of 748 bp in length
*    34368 34467: gap of      100 bp
*    34468    35194: contig of 727 bp in length
*    35195 35294: gap of      100 bp
*    35295    36021: contig of 727 bp in length
*    36022 36121: gap of      100 bp
*    36122    36834: contig of 713 bp in length
*    36835 36934: gap of      100 bp
*    36935    37649: contig of 715 bp in length
*    37650 37749: gap of      100 bp
*    37750    38466: contig of 717 bp in length
*    38467 38566: gap of      100 bp
*    38567    39289: contig of 723 bp in length
*    39290 39389: gap of      100 bp
*    39390    40121: contig of 732 bp in length
*    40122 40221: gap of      100 bp
*    40222    40950: contig of 729 bp in length
*    40951 41050: gap of      100 bp
*    41051    41777: contig of 727 bp in length
*    41778 41877: gap of      100 bp
*    41878    42617: contig of 740 bp in length
*    42618 42717: gap of      100 bp
*    42718    43431: contig of 714 bp in length
*    43432 43531: gap of      100 bp
*    43532    44252: contig of 721 bp in length
*    44253 44352: gap of      100 bp
*    44353    45064: contig of 712 bp in length
*    45065 45164: gap of      100 bp
*    45165    45872: contig of 708 bp in length
*    45873 45972: gap of      100 bp
*    45973    46667: contig of 695 bp in length
*    46668 46767: gap of      100 bp
*    46768    47497: contig of 730 bp in length
*    47498 47597: gap of      100 bp
*    47598    48326: contig of 729 bp in length
*    48327 48426: gap of      100 bp
*    48427    49164: contig of 738 bp in length
*    49165 49264: gap of      100 bp
*    49265    49988: contig of 724 bp in length
*    49989 50088: gap of      100 bp
*    50089    50813: contig of 725 bp in length
*    50814 50913: gap of      100 bp
*    50914    51645: contig of 732 bp in length
*    51646 51745: gap of      100 bp
*    51746    52462: contig of 717 bp in length
*    52463 52562: gap of      100 bp
*    52563    53280: contig of 718 bp in length
*    53281 53380: gap of      100 bp
*    53381    54092: contig of 712 bp in length
*    54093 54192: gap of      100 bp
*    54193    54908: contig of 716 bp in length
*    54909 55008: gap of      100 bp
*    55009    55733: contig of 725 bp in length
*    55734 55833: gap of      100 bp
*    55834    56569: contig of 736 bp in length
*    56570 56669: gap of      100 bp

  Query Match             3.6%;  Score 43.4;  DB 77;  Length 62336;
  Best Local Similarity   34.9%;  Pred. No. 0.12;
  Matches  149;  Conservative    0;  Mismatches  278;  Indels    0;  Gaps    0;

Qy    633 TGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGC 692
          | |||  |   |   | || ||| || ||    ||        |    |  ||   |||
Db  59214 TTGCCTAAAAGGGAAGCGTTTTTTCCNACCCCACAATGGAAGGGTTANATACCAAGTTGC 59155

Qy    693 CAACAGCTTAAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCC 752
```

```
Db  59154 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 59095

Qy    753 AGACATCTACCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTG 812
                                                 || || || ||||| | ||
Db  59094 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTTTTTTACGTTTTTATTT 59035

Qy    813 CGTGAAGGGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCT 872
          ||| |   |   ||      |  ||||| ||| | |||||  ||  |||  |
Db  59034 TGTGTTGAAGACAAACGGAGGGAAAAAAGACGAAAAGGAAAAACCAGCGACAAGGGGGGG 58975

Qy    873 GGCTAGAGAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGG 932
          ||   |  |  | ||  |||||||| |      |     |     |     || || ||
Db  58974 GGGGGGGCAGGGAAGGGGGAGAAGAGGCGCCGGGGAGAAAAAAGGCGAAACAGGGAGAGG 58915

Qy    933 GCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAAC 992
          |   |  |   ||   |  || ||| ||   | || ||||   || |      || |||
Db  58914 GGGCGAGGGAACGAACCAGGAAGGGGAGACAGAGGGGGCAGAAAGGAGAGAGAAGGAAAG 58855

Qy    993 CCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAA 1052
          | |||||   | |      | |  ||     |  |      |  || | || ||| ||
Db  58854 CAAGAGCGGGGGGAGCGACAAACAAAGAGGAAGAGAACGGGGAGAGAGGAGAAGAAGGAA 58795

Qy   1053 AGGGGAA 1059
          ||  ||
Db  58794 AGAAAAA 58788


RESULT  15
CNS0763N/c
LOCUS       CNS0763N     1111 bp    DNA             STS       11-JAN-2001
DEFINITION  clone XBA0AB001F07 of library XBA0AB from strain CLIB 210 of
            Kluyveromyces lactis, sequence tagged site.
ACCESSION   AL430905
VERSION     AL430905.1  GI:12214099
KEYWORDS    STS.
SOURCE      Kluyveromyces lactis.
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 1111)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
   PUBMED   11152886
REFERENCE   2  (bases 1 to 1111)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
            yeast species for molecular evolution studies(1)
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   PUBMED   11152876
REFERENCE   3  (bases 1 to 1111)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This STS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .1111
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="XBA0AB001F07"
                     /clone_lib="XBA0AB"
     STS             1. .1111
                     /note="part of mitochondrial DNA"
                     /evidence=not_experimental
BASE COUNT      173 a    366 c     58 g    416 t     98 others
ORIGIN


  Query Match             3.5%;  Score 41.8;  DB 53;  Length 1111;
  Best Local Similarity   42.6%;  Pred. No. 0.23;
  Matches  100;  Conservative   16;  Mismatches  119;  Indels    0;  Gaps    0;

Qy    823 AAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAA 882
          |||:     |  |   |||| |||||||:| |||    | |  |:: |:  :|  :| |
Db    918 AAGRGGGGAAAAAAAANAAGAAAAAAGARAAAAAAAAAAAAAAGARRGARGNRGGGRGGGGG 859

Qy    883 AGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCT 942
          :|| |   :| || || || :: |    | |           ||  | |||   |  |
Db    858 RGGGGRRARAAAAAAAAAAARRRGGGGGGGGGGGGKNNNNGGGGGGGGGGGGGGGGGGGGG 799

Qy    943 GCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATC 1002
          | | ||   | ||||  ||||| || ||    | :     |  :|: :|   | |  |
Db    798 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGRGGGGGGARGRGRAAAAAAAAAAAA 739

Qy   1003 AGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGG 1057
          | |||| | | |    ||    ||  |      | || ||||| || |  :||||
Db    738 AAAAAAGGGGGGGRGGGGGARAGGGRAAAGGAAGAAAAAAAAAAAAAAKGRGGGG 684


Search completed: June 21, 2001, 02:12:08
Job time: 8597 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2001, 02:08:54 ; Search time 153.2 Seconds
(without alignments)
4584.131 Million cell updates/sec

Title: US-09-558-474-3
Perfect score: 1203
Sequence: 1 CGCTTAGAAGTCGGACTACA.....TACCAATACGAACTGACAAA 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0401:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1203	100.0	1203	20	Z08866	Mouse interleukin
	2	1203	100.0	1203	20	X17787	Mouse interleukin-
	3	1203	100.0	1203	21	A52578	Murine interleukin
	4	439.8	36.6	1026	21	Z37262	SGRF coding sequen
	5	439.8	36.6	1055	22	C84306	Human EXCS encodin
	6	439.8	36.6	1067	21	A16686	Human secreted pro
	7	391	32.5	570	20	Z08865	Human interleukin
	8	391	32.5	570	20	X17786	Human interleukin-
	9	391	32.5	570	21	A52577	Human interleukin-
	10	203.2	16.9	412	21	A42676	Human secreted exp
	11	151.6	12.6	2398	21	Z37263	SGRF coding sequen
c	12	35.4	2.9	9312	21	Z43768	Human CRP-ductin c
	13	35	2.9	396	21	C06417	Human secreted pro
	14	34.8	2.9	292	21	C09932	Human secreted pro
	15	34.6	2.9	3960	20	X83478	A.thaliana potassi
c	16	34.4	2.9	3492	20	X35745	cDNA encoding a pr.
	17	34.4	2.9	10732	21	A10594	Gene encoding a su
	18	33.8	2.8	729	22	F44973	Murine TANGO 253 c
	19	33.8	2.8	729	22	F45038	Murine secreted pr
	20	33.8	2.8	729	22	F45039	Murine secreted pr
	21	33.8	2.8	729	22	F45040	Murine secreted pr
	22	33.8	2.8	729	22	F45041	Murine secreted pr
	23	33.8	2.8	1263	22	F44972	Murine TANGO 253 O
	24	33.8	2.8	1263	22	F45006	Murine secreted pr
	25	33.8	2.8	1263	22	F45007	Murine secreted pr
	26	33.8	2.8	1263	22	F45008	Murine secreted pr
	27	33.8	2.8	1263	22	F45009	Murine secreted pr
	28	33.6	2.8	2620	20	Z22302	cDNA encoding a hu
	29	33.2	2.8	282	21	C00385	Human secreted pro
	30	33.2	2.8	1038602	20	Z01425	Complete genome se
c	31	33	2.7	267	16	T22945	Human gene signatu
c	32	33	2.7	411	18	T73898	Humanised variable
	33	33	2.7	1514	21	C38959	Arabidopsis thalia
c	34	33	2.7	3963	21	Z52428	HTRM clone 2724537
c	35	32.6	2.7	1505	15	Q55750	Genomic clone G11F
	36	32.6	2.7	50000	21	A64140	Nucleotide sequenc
	37	32.4	2.7	482	21	C37809	Arabidopsis thalia
	38	32.4	2.7	2324	21	A48578	cDNA encoding whea
	39	32.4	2.7	8967	20	X83938	Salmonella typhimu
c	40	32.4	2.7	8967	20	X83940	Salmonella typhimu
	41	32.4	2.7	8967	20	X83942	Salmonella typhimu
c	42	32.4	2.7	8967	20	X83944	Salmonella typhimu
	43	32.4	2.7	24701	20	X83935	Salmonella typhimu
c	44	32.4	2.7	24701	20	X83936	Salmonella typhimu
c	45	32.2	2.7	1700	20	Z31921	Human helicase, Re

ALIGNMENTS

RESULT 1
Z08866
ID Z08866 standard; cDNA; 1203 BP.
XX
AC Z08866;
XX
DT 04-NOV-1999 (first entry)
XX
DE Mouse interleukin B30 encoding cDNA.
XX
KW DNAX soluble receptor subunit 1; DNAX cytokine receptor subunit 1;
KW interleukin B30; DSRS1; DCRS1; IL-B30; cytokine receptor; diagnosis;
KW inflammatory disorder; inflammatory response; innate immunity;
KW morphogenic development; immunological disorder; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 113..703
FT /*tag= a
FT /product= "IL-B30"
FT sig_peptide 113..175
FT /*tag= b
FT mat_peptide 176..700
FT /*tag= c
XX
PN WO9940195-A1.
XX
PD 12-AUG-1999.
XX
PF 05-FEB-1999; 99WO-US02600.
XX
PR 13-MAY-1998; 98US-0078194.
PR 06-FEB-1998; 98US-0073941.

```
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Kastelein RA,  Mattson JD,  McClanahan TK;
XX
DR   WPI; 1999-527306/44.
DR   P-PSDB; Y29784.
XX
PT   New receptor subunits useful in the treatment inflammatory disorders
XX
PS   Claim 16; Page 27-28; 133pp; English.
XX
CC   The present invention describes a composition (I) comprising DNAX
CC   cytokine receptor subunit I (DCRS1) protein and DNAX soluble receptor
CC   subunit I (DSRS1) protein, which together encode a new mammalian
CC   cytokine-related receptor (R), or DCRS1 and interleukin B30 (IL-B30)
CC   proteins, or DSRS1 and IL-B30 proteins. (I) comprising DSRS1 and DCRS1
CC   is useful for screening for ligands (i.e. agonists/antagonists) from
CC   a library of compounds, which are useful for modulating the physiology
CC   or development of a cell or tissue culture e.g. inflammatory responses,
CC   innate immunity and/or morphogenic development. (R), antibodies and
CC   ligands are useful for treatment of conditions, especially immunological
CC   disorders, associated with conditions exhibiting abnormal expression of
CC   (R). (R) is useful as a phosphate labeling enzyme to label substrates,
CC   and the subunits DSRS1 and DCRS1 are useful as immunogens for generating
CC   antibodies, or as antigens for binding antibodies. Nucleic acids
CC   encoding (R) are useful for identifying related DNAs and mRNAs, and
CC   variants from other individuals or species. The present sequence
CC   encodes the specifically claimed mouse IL-B30, for use in the
CC   composition of the present invention.
XX
SQ   Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;

  Query Match             100.0%;  Score 1203;  DB 20;  Length 1203;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cgcttagaagtcggactacagagttagactcagaaccaaaggaggtggataggggtcca 60

Qy      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 caggcctggtgcagatcacagagccagccagatctgagaagcagggaacaagatgctgga 120

Qy     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ttgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctggctgtgcc 180

Qy     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 taggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctctgcatgct 240

Qy     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 agcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaagaggatga 300

Qy     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 agagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgacccacaaggact 360

Qy     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 caaggacaacagccagttctgcttgcaaaggatccgccaaggtctggctttttataagca 420

Qy     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagccccatgga 480

Qy     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 gcaacttcacacctccctactaggactcagccaactcctccagccagaggatcaccccg 540

Qy     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgccccttctccg 600

Qy     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtctttgccca 660

Qy     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 cggagcagcaactctgactgagcccttagtgccaacagcttaaggatgcccaggttccca 720

Qy     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggctaccatgataagactaatctatcagcccagacatctaccagttaattaacccatta 780

Qy     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ggacttgtgctgttcttgtttcgtttgttttgcgtgaagggcaaggacaccattattaaa 840

Qy     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gagaaaagaaacaaaccccagagcaggcagctggctagagaaaggagctggagaagaaga 900

Qy     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ataaagtctcgagcccttggccttggaagcgggcaagcagctgcgtggcctgaggggaag 960

Qy     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ggggcggtggcatcgagaaactgtgagaaaacccagagcatcagaaaaagtgagcccagg 1020

Qy    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ctttggccattatctgtaagaaaaacaagaaaaggggaacattatactttcctgggtggc 1080

Qy    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 tcagggaaatgtgcagatgcacagtactccagacagcagctctgtacctgcctgctctgt 1140

Qy    1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ccctcagttctaacagaatctagtcactaagaactaacaggactaccaatacgaactgac 1200

Qy    1201 AAA 1203
          |||
Db    1201 aaa 1203


RESULT   2
X17787
ID   X17787 standard; cDNA; 1203 BP.
XX
AC   X17787;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Mouse interleukin-B30 (IL-B30) polypeptide encoding cDNA.
XX
KW   Cytokine; interleukin-B30; IL-B30; forensic science; cell proliferation;
KW   inflammatory condition; drug screening; mouse; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..703
FT                   /*tag= a
FT                   /product= "IL-B30"
FT   sig_peptide     113..175
FT                   /*tag= b
FT   mat_peptide     176..700
FT                   /*tag= c
```

XX
PN WO9905280-A1.
XX
PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-US15423.
XX
PR 25-JUL-1997; 97US-0900905.
XX
PA (SCHE) SCHERING CORP.
XX
PI Bazan JF;
XX
DR WPI; 1999-142935/12.
DR P-PSDB; W95003.
XX
PT Newly isolated or recombinant polynucleotide encoding mammalian
PT cytokine interleukin-B30 (IL-B30), including fragments - useful for
PT regulating activation, development, differentiation and function of
PT various cell types, and for diagnosing and treating conditions
PT associated with IL-B30
XX
PS Claim 3; Page 9-10; 83pp; English.
XX
CC This cDNA encodes a mouse cytokine interleukin-B30 (IL-B30) polypeptide.
CC Host cells containing a vector comprising the IL-B30 nucleic acid are
CC used for the recombinant production of the protein. The polynucleotides
CC are useful for diagnosis of IL-B30 mediated conditions, and forensic
CC science (e.g. to distinguish rodent from human, or as a marker to
CC distinguish between different cells exhibiting differential expression or
CC modification patterns). The IL-B30 (including fragments), together with
CC antibodies that bind to IL-B30 are useful for teaching purposes. They
CC are also used for treating conditions associated with abnormal physiology
CC or development, including inflammatory conditions. The polypeptide
CC cytokine should mediate cytokine synthesis and proliferation in cells.
CC IL-B30 is useful for drug screening to identify compounds having binding
CC affinity to IL-B30.
XX
SQ Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;

Query Match 100.0%; Score 1203; DB 20; Length 1203;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cgcttagaagtcggactacagagttagactcagaaccaaaggaggtggataggggtccca 60

Qy      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 caggcctggtgcagatcacagagccagccagatctgagaagcagggaacaagatgctgga 120

Qy     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ttgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctggctgtgcc 180

Qy     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 taggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctctgcatgct 240

Qy     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 agcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaagaggatga 300

Qy     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 agagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgacccacaaggact 360

Qy     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 caaggacaacagccagttctgcttgcaaaggatccgccaaggtctggctttttataagca 420

Qy     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagccccatgga 480

Qy     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 gcaacttcacacctccctactaggactcagccaactcctccagccagaggatcaccccg 540

Qy     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgcccccttctccg 600

Qy     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtctttgccca 660

Qy     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 cggagcagcaactctgactgagcccttagtgccaacagcttaaggatgcccaggttccca 720

Qy     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggctaccatgataagactaatctatcagcccagacatctaccagttaattaacccatta 780

Qy     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ggacttgtgctgttcttgtttcgtttgttttgcgtgaagggcaaggacaccattattaaa 840

Qy     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gagaaaagaaacaaaccccagagcaggcagctggctagagaaaggagctggagaagaaga 900

Qy     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ataaagtctcgagcccttggccttggaagcgggcaagcagctgcgtggcctgaggggaag 960

Qy     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ggggcggtggcatcgagaaactgtgagaaaacccagagcatcagaaaaagtgagcccagg 1020

Qy    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ctttggccattatctgtaagaaaaacaagaaaaggggaacattatactttcctgggtggc 1080

Qy    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 tcagggaaatgtgcagatgcacagtactccagacagcagctctgtacctgcctgctctgt 1140

Qy    1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ccctcagttctaacagaatctagtcactaagaactaacaggactaccaatacgaactgac 1200

Qy    1201 AAA 1203
           |||
Db    1201 aaa 1203
```

RESULT 3
A52578
ID A52578 standard; cDNA; 1203 BP.
XX
AC A52578;
XX
DT 27-SEP-2000 (first entry)
XX
DE Murine interleukin-B30 (IL-B30) cDNA.
XX
KW Interleukin-B30; IL-B30; murine; mouse; cytokine; cellular signalling;
KW immune response; haematopoietic cell; IL-6 homologue; GCSF homologue;
KW granulocyte colony stimulating factor; immune disorder;
KW inflammatory disease; autoimmune disease; antigen; antibody; ss.
XX
OS Mus sp.

```
XX
FH   Key             Location/Qualifiers
FT   CDS             113..703
FT                   /*tag=  a
FT                   /product= "Murine IL-B30 precursor"
FT   sig_peptide     113..175
FT                   /*tag=  b
FT   mat_peptide     176..700
FT                   /*tag=  c
FT                   /product= "Mature murine IL-B30"
XX
PN   US6060284-A.
XX
PD   09-MAY-2000.
XX
PF   24-JUL-1998;   98US-0122443.
XX
PR   25-JUL-1997;   97US-0053765.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 2000-364420/31.
DR   P-PSDB; B01982.
XX
PT   Novel recombinant DNA encoding cytokines especially interleukin-B30
PT   useful as probes or primers for diagnosing immune disorders including
PT   autoimmune or chronic inflammatory conditions -
XX
PS   Claim 3; Column 7-10; 32pp; English.
XX
CC   This sequence represents cDNA encoding murine interleukin-B30 (IL-B30).
CC   IL-B30 is a novel cytokine, exhibiting significant homology to IL-6 and
CC   GCSF (granulocyte colony stimulating factor). Cytokines play a critical
CC   role in signalling between immune or other cells during an immune
CC   response. The precise role if IL-B30 is not yet known - it is likely to
CC   have either a stimulatory or an inhibitory effect on haematopoietic cells
CC   such as T-cells, B-cells, natural killer (NK) cells and macrophages.
CC   Alternatively, it may affect vascular physiology or development, or have
CC   neuronal effects. IL-B30, its fragments, IL-B30 nucleotides, agonists and
CC   antagonists are useful in the diagnosis and treatment of disorders
CC   associated with abnormal expression or activity of IL-B30 e.g.,
CC   inflammatory or autoimmune disorders. Nucleotides encoding IL-B30 are
CC   useful for recombinant expression of IL-B30 in a host cell, and as a
CC   source of probes and primers. The IL-B30 probes and primers can be used
CC   to detect levels of IL-B30 expression in samples from patients suspected
CC   of having an inflammatory or autoimmune disorder. IL-B30 nucleotides may
CC   also be used to identify homologous genes in other species. IL-B30
CC   protein or its fragments are useful as antigens for raising antibodies to
CC   various linear and conformational epitopes. Such antibodies may be used
CC   to detect levels of IL-B30 protein in a sample.
XX
SQ   Sequence 1203 BP; 336 A; 310 C; 312 G; 245 T; 0 other;

  Query Match             100.0%;  Score 1203;  DB 21;  Length 1203;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cgcttagaagtcggactacagagttagactcagaaccaaaggaggtggataggggtcca 60

Qy      61 CAGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 caggcctggtgcagatcacagagccagccagatctgagaagcagggaacaagatgctgga 120

Qy     121 TTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ttgcagagcagtaataatgctatggctgttgccctgggtcactcagggcctggctgtgcc 180

Qy     181 TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 taggagtagcagtcctgactgggctcagtgccagcagctctctcggaatctctgcatgct 240

Qy     241 AGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 agcctggaacgcacatgcaccagcgggacatatgaatctactaagagaagaagaggatga 300

Qy     301 AGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 agagactaaaaataatgtgccccgtatccagtgtgaagatggttgtgacccacaaggact 360

Qy     361 CAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 caaggacaacagccagttctgcttgcaaaggatccgccaaggtctggcttttttataagca 420

Qy     421 CCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cctgcttgactctgacatcttcaaaggggagcctgctctactccctgatagccccatgga 480

Qy     481 GCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 gcaacttcacacctccctactaggactcagccaactcctccagccagaggatcacccccg 540

Qy     541 GGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ggagacccaacagatgcccagcctgagttctagtcagcagtggcagcgcccccttctccg 600

Qy     601 TTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ttccaagatccttcgaagcctccaggccttttggccatagctgcccgggtctttgcccca 660

Qy     661 CGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 cggagcagcaactctgactgagcccttagtgccaacagcttaaggatgcccaggttccca 720

Qy     721 TGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggctaccatgataagactaatctatcagcccagacatctaccagttaattaacccatta 780

Qy     781 GGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATTAAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ggacttgtgctgttcttgtttcgtttgttttgcgtgaagggcaaggacaccattattaaa 840

Qy     841 GAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gagaaaagaaacaaaccccagagcaggcagctggctagagaaaggagctggagaagaaga 900

Qy     901 ATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ataaagtctcgagcccttggccttggaagcgggcaagcagctgcgtggcctgaggggaag 960

Qy     961 GGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ggggcggtggcatcgagaaactgtgagaaaacccagagcatcagaaaaagtgagcccagg 1020

Qy    1021 CTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ctttggccattatctgtaagaaaaacaagaaaaggggaacattatactttcctgggtggc 1080

Qy    1081 TCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 tcagggaaatgtgcagatgcacagtactccagacagcagctctgtacctgcctgctctgt 1140

Qy    1141 CCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ccctcagttctaacagaatctagtcactaagaactaacaggactaccaatacgaactgac 1200

Qy    1201 AAA 1203
           |||
Db    1201 aaa 1203

RESULT   4
```

```
Z37262
ID   Z37262 standard; DNA; 1026 BP.
XX
AC   Z37262;
XX
DT   01-FEB-2000  (first entry)
XX
DE   SGRF coding sequence.
XX
KW   SGRF; human; Interleukin-6 G-CSF related factor; cell proliferation;
KW   immune system; haematopoietic system; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9954357-A1.
XX
PD   28-OCT-1999.
XX
PF   14-APR-1999;   99WO-JP01997.
XX
PR   14-APR-1998;   98JP-0121805.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Hirata Y;
XX
DR   WPI; 2000-013230/01.
DR   P-PSDB; Y54606.
XX
PT   Novel cytokine-like protein, with activity of supporting proliferation
PT   of myeloid cells, useful in treating abnormality of cell proliferation
PT   in immune and haematopoiesis systems -
XX
PS   Claim 2; Fig 1; 69pp; Japanese.
XX
CC   This sequence encodes the Interleukin-6 G-CSF related factor (SGRF)
CC   protein of the invention. The protein is a member of the IL-6/G-CSF/MSF
CC   family. The protein can be used in drugs for treating diseases due to
CC   abnormality of cell proliferation in the immune system and haematopoietic
CC   system.
XX
SQ   Sequence 1026 BP; 268 A; 249 C; 272 G; 237 T; 0 other;


  Query Match             36.6%;  Score 439.8;  DB 21;  Length 1026;
  Best Local Similarity   77.4%;  Pred. No. 1.9e-123;
  Matches  600;  Conservative    0;  Mismatches  162;  Indels   13;  Gaps    5;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      48 agagaatcaggctcaaagcaagtggaagtgggcagagattccaccaggactggtgcaagg 107

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||   ||||||| |||||
Db     108 cgcagagccagccagatttgagaagaaggcaaaaagatgctggggagcagagctgtaatg 167

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
            |||| | |||| ||||||||    ||||||||   |||||||| || | ||||| |||
Db     168 ctgctgttgctgctgccctggacagctcagggcagagctgtgcctgggggcagcagccct 227

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     228 gcctggactcagtgccagcagctttcacagaagctctgcacactggcctggagtgcacat 287

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
            ||| || |||||| ||| ||   |||||||||| |  |||||||||||| |||| ||
Db     288 ccactagtgggacacatggat---ctaagagaagagggagatgaagagactacaaatgat 344

Qy     317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
           || |||| |||||||||| |||||| |||||||| |||||||||| ||||||||| |||
Db     345 gttccccatatccagtgtggagatggctgtgacccccaaggactcagggacaacagtcag 404

Qy     377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
           |||||||||||||||||| ||| ||||||  ||||||| || | ||||| |  || ||
Db     405 ttctgcttgcaaaggatccaccagggtctgatttttttatgagaagctgctaggatcggat 464

Qy     437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
           || |||| |||||||||| |||| ||||||||||||||  |||  || |||||  |||||
Db     465 attttcacaggggagccttctctgctccctgatagccctgtgggccagcttcatgcctcc 524

Qy     497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
           ||||| || |||||||||||||| ||||| |||| ||||| | |||||||| || |||||
Db     525 ctactgggcctcagccaactcctgcagcctgagggtcaccactgggagactcagcagatt 584

Qy     557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
           || ||||| ||| | || |||| |||||||| | ||||||||| | ||| ||||||||
Db     585 ccaagcctcagtcccagccagccatggcagcgtctccttctccgcttcaaaatccttcgc 644

Qy     617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
           |||||||||||||| |||| |||| ||||||||||||||||| ||||||||||
Db     645 agcctccaggcctttgtggctgtagccgcccgggtctttgcccatggagcagcaac---- 700

Qy     677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
           ||||| || ||  | || |||| |||||||   | ||| |  |||||  |  | | |
Db     701 cctgagtccctaaaggcagcagctcaaggatggcactcagatctccatggcccagcaagg 760

Qy     732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
             ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db     761 ccaagataaatctaccaccccaggcacctgtgagccaacaggttaattagtccat 815


RESULT   5
C84306
ID   C84306 standard; cDNA; 1055 BP.
XX
AC   C84306;
XX
DT   19-MAR-2001  (first entry)
XX
DE   Human EXCS encoding cDNA (clone ID 2933038CB1).
XX
KW   Extracellular signaling molecule; EXCS; anti-inflammatory; human;
KW   immunosuppressive; cytostatic; neuroprotective; gastrointestinal;
KW   virucide; antibacterial; anti-HIV; human immunodeficiency virus;
KW   antiinfertility; cerebroprotective; nootropic; antiulcer; antifungal;
KW   anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological;
KW   keratolytic; protozoacide; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200070049-A2.
XX
PD   23-NOV-2000.
XX
PF   19-MAY-2000; 2000WO-US13975.
XX
PR   19-MAY-1999;   99US-0134949.
PR   15-JUL-1999;   99US-0144270.
PR   30-JUL-1999;   99US-0146700.
PR   04-OCT-1999;   99US-0157508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Lal P,  Burford N,  Bandman O,  Baughn MR;
PI   Azimzai Y,  Lu DAM,  Patterson C;
XX
DR   WPI; 2001-025021/03.
DR   P-PSDB; B48070.
XX
PT   New human extracellular signaling nucleic acids and polypeptides useful
PT   for diagnosing, treating and preventing infections and
PT   gastrointestinal, neurological, reproductive, and
PT   autoimmune/inflammatory disorders -
XX
PS   Claim 4; Page 106; 114pp; English.
XX
```

```
CC   The invention provides human extracellular signaling molecules (EXCS)
CC   and polynucleotides which identify and encode EXCS. EXCS can be expressed
CC   by standard recombinant methodology. The amino acid and nucleic acid
CC   sequences of EXCS are useful for diagnosing, treating and preventing
CC   infections and gastrointestinal (peptic ulcer, dysphagia, pancreatitis),
CC   neurological (e.g. epilepsy, ischemic cerebrovascular disease, stroke),
CC   reproductive (infertility, ovulatory defects, endometriosis), autoimmune
CC   /inflammatory (actinic keratosis, acquired immunodeficiency syndrome
CC   (AIDS), Addison's disease), and cell proliferative disorders including
CC   cancers (of the breast, adrenal gland, bone). They may also be used to
CC   treat fatal familial insomnia, nutritional and metabolic diseases of the
CC   nervous system, myopathies, mental disorders (anxiety, schizophrenia,
CC   mood), as well as infections caused by parasites (malaria, leishmania,
CC   trypanosoma), viral (adenovirus, coronavirus, flavivirus), bacterial
CC   (e.g. pneumococcus, staphylococcus, bacillus), and fungal (aspergillus,
CC   blastomyces, dermatophytes) agents. The nucleic acids, polypeptides,
CC   antagonists, agonists, pharmaceutical compositions, and antibodies may
CC   also be used for treating or preventing disorders associated with
CC   increased or decreased expression or activity of EXCS. EXCS
CC   polynucleotides may also be used to detect and quantify gene expression
CC   in biopsied tissues in which expression of EXCS may be correlated with
CC   the disease, to determine presence or excess expression of EXCS, to
CC   monitor regulation of EXCS levels during therapeutic intervention, to
CC   detect the presence of associated disorders, as targets in microarray, to
CC   generate hybridization probes, and to detect differences in gene
CC   sequences among normal, carrier or affected individuals. Antibodies may
CC   also be used in diagnosing disorders, in monitoring patients being
CC   treated with EXCS agonists, antagonists or inhibitors. Sequences
CC   C84293-C84318 represent nucleic acid molecules encoding the EXCS of the
CC   invention.
XX
SQ   Sequence 1055 BP; 284 A; 254 C; 278 G; 239 T; 0 other;

  Query Match             36.6%;  Score 439.8;  DB 22;  Length 1055;
  Best Local Similarity   77.4%;  Pred. No. 1.9e-123;
  Matches  600;  Conservative    0;  Mismatches  162;  Indels   13;  Gaps    5;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      75 agagaatcaggctcaaagcaagtggaagtgggcagagattccaccaggactggtgcaagg 134

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || |||||||||   ||||||| |||||
Db     135 cgcagagccagccagatttgagaagaaggcaaaaagatgctggggagcagagctgtaatg 194

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
            |||| | |||| ||||||||    ||||||||   |||||||| || | ||||| |||
Db     195 ctgctgttgctgctgccctggacagctcagggcagagctgtgcctgggggcagcagccct 254

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     255 gcctggactcagtgccagcagctttcacagaagctctgcacactggcctggagtgcacat 314

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
           ||| || |||||| ||| ||   |||||||||| |  |||||||||||| |||| ||
Db     315 ccactagtgggacacatggat---ctaagagaagagggagatgaagagactacaaatgat 371

Qy     317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
           || |||| |||||||||| |||||| |||||||| |||||||||| ||||||||| |||
Db     372 gttccccatatccagtgtggagatggctgtgacccccaaggactcagggacaacagtcag 431

Qy     377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
           |||||||||||||||||| ||| ||||||  ||||||| || | ||||| |  || ||
Db     432 ttctgcttgcaaaggatccaccagggtctgatttttttatgagaagctgctaggatcggat 491

Qy     437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
           || |||| |||||||||| |||| ||||||||||||||  |||  || |||||  |||||
Db     492 attttcacaggggagccttctctgctccctgatagccctgtgggccagcttcatgcctcc 551

Qy     497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
           ||||| || |||||||||||||| ||||| |||| ||||| | ||||||| || |||||
Db     552 ctactgggcctcagccaactcctgcagcctgagggtcaccactgggagactcagcagatt 611

Qy     557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
           || ||||| ||| | || |||| |||||||| | ||||||||| | ||| ||||||||
Db     612 ccaagcctcagtcccagccagccatggcagcgtctccttctccgcttcaaaatccttcgc 671

Qy     617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
           |||||||||||||||| ||||  |||| |||||||||||||||| ||||||||||
Db     672 agcctccaggcctttgtggctgtagccgcccgggtctttgcccatggagcagcaac---- 727

Qy     677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
           ||||| || ||  | || ||||| |||||||   | ||| |  ||||||  |  | | |
Db     728 cctgagtccctaaaggcagcagctcaaggatggcactcagatctccatggcccagcaagg 787

Qy     732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
             ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db     788 ccaagataaatctaccaccccaggcacctgtgagccaacaggttaattagtccat 842


RESULT   6
A16686
ID   A16686 standard; cDNA; 1067 BP.
XX
AC   A16686;
XX
DT   16-JUN-2000  (first entry)
XX
DE   Human secreted protein clone rk80_3 nucleotide sequence SEQ ID NO:137.
XX
KW   Human; secreted protein; immunestimulant; immunesuppressant; virucide;
KW   antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
KW   antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;
KW   antithyroid; immune deficiency; severe combined immunedeficiency; SCID;
KW   infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
KW   connective tissue disease; multiple sclerosis; erythematosis;
KW   rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
KW   Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
KW   insulin dependent diabetes mellitus; graft-versus-host-disease;
KW   autoimmune inflammatory eye disease; allergy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009552-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1999;   99WO-US18298.
XX
PR   14-AUG-1998;   98US-0096622.
PR   17-AUG-1998;   98US-0096815.
PR   04-SEP-1998;   98US-0099229.
PR   23-OCT-1998;   98US-0105368.
PR   08-JAN-1999;   99US-0115234.
PR   12-FEB-1999;   99US-0119931.
PR   18-FEB-1999;   99US-0120575.
PR   30-APR-1999;   99US-0132020.
PR   11-AUG-1999;   99US-0096622.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M,  Agostino MJ,  Steininger RJ,  Spaulding V;
PI   Wong GG,  Clark HF,  Fechtel K;
XX
DR   WPI; 2000-205979/18.
DR   P-PSDB; Y94966.
XX
PT   New polynucleotides encoding secreted proteins, which may have e.g.
PT   nutritional, chemokine, immune stimulating or suppressing,
PT   hematopoiesis regulating, tissue growth, activin/inhibin
PT   antiinflammatory or tumor inhibition activity  -
XX
PS   Claim 146; Page 596-597; 641pp; English.
XX
```

```
CC   A16618 to A16697 encode the human secreted proteins given in Y94898
CC   to Y94980, isolated from human adult brain, adult thyroid, adult retina,
CC   foetal carcinoma, adult blood, adult neural, foetal kidney, adult
CC   placenta, adult testis, whole embryo, adult cartilage, kidney, foetal
CC   brain, adult thymus, foetal placenta, adult uterus, adult tumour, and
CC   adult bladder, cDNA libraries. The polynucleotides and proteins are
CC   predicted to have biological activities which would make them suitable
CC   for treating, preventing or ameliorating medical conditions in humans
CC   and animals. The polynucleotides can be used as markers for tissues in
CC   which the protein is preferentially expressed, as molecular weight
CC   markers on Southern gels, and as chromosome markers or tags to identify
CC   chromosomes or to map gene positions. The proteins can be used in the
CC   treatment of immune deficiencies and disorders, such as severe combined
CC   immunedeficiency (SCID), as well as viral, bacterial, fungal and other
CC   infections. These infections include human immunodeficiency virus (HIV),
CC   hepatitis, herpesviruses, mycobacteria, Leismania spp., malaria and
CC   candidiasis. The proteins can be used to treat autoimmune disorders such
CC   as connective tissue disease, multiple sclerosis, systemic lupus
CC   erythematosis, rheumatoid arthritis, autoimmune pulmonary inflammation,
CC   Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
CC   diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
CC   autoimmune inflammatory eye disease. The proteins can also be used to
CC   treat allergic conditions, such asthma. A16698 to A16774 represent
CC   probes for the human secreted proteins from the present invention.
XX
SQ   Sequence 1067 BP; 293 A; 255 C; 281 G; 238 T; 0 other;


  Query Match             36.6%;  Score 439.8;  DB 21;  Length 1067;
  Best Local Similarity   77.4%;  Pred. No. 1.9e-123;
  Matches  600;  Conservative    0;  Mismatches  162;  Indels   13;  Gaps    5;

Qy     18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
          | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db     83 agagaatcaggctcaaagcaagtggaagtgggcagagattccaccaggactggtgcaagg 142

Qy     77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
          | ||||||||||||||| |||||| ||| || |||||||||   |||||| ||||
Db    143 cgcagagccagccagatttgagaagaaggcaaaaagatgctggggagcagagctgtaatg 202

Qy    137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
           |||| | |||| |||||||||    ||||||||   |||||||| || | ||||| |||
Db    203 ctgctgttgctgctgccctggacagctcagggcagagctgtgcctgggggcagcagccct 262

Qy    197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
          | |||| ||||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db    263 gcctggactcagtgccagcagctttcacagaagctctgcacactggcctggagtgcacat 322

Qy    257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
           ||| || |||||| ||| ||   |||||||||| |  ||||||||||||| |||| ||
Db    323 ccactagtgggacacatggat---ctaagagaagagggagatgaagagactacaaatgat 379

Qy    317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAG 376
          || |||| |||||||||| |||||| |||||||| |||||||||| ||||||||| |||
Db    380 gttccccatatccagtgtggagatggctgtgacccccaaggactcagggacaacagtcag 439

Qy    377 TTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGAC 436
          ||||||||||||||||||| ||| |||||| |||||| || | ||||| |  || ||
Db    440 ttctgcttgcaaaggatccaccagggtctgatttttatgagaagctgctaggatcggat 499

Qy    437 ATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCC 496
          || |||| |||||||||| |||| |||||||||||||  ||| || ||||| |||||
Db    500 attttcacaggggagccttctctgctccctgatagccctgtgggccagcttcatgcctcc 559

Qy    497 CTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATG 556
          ||||| || |||||||||||||| ||||| |||| ||||| | ||||||| || |||||
Db    560 ctactgggcctcagccaactcctgcagcctgagggtcaccactgggagactcagcagatt 619

Qy    557 CCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGA 616
          || ||||| ||| | || ||||  |||||||| | ||||||||| | ||| ||||||||
Db    620 ccaagcctcagtcccagccagccatggcagcgtctccttctccgcttcaaaatccttcgc 679

Qy    617 AGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTG 676
          |||||||||||||||| ||||  |||| |||||||||||||||| ||||||||||
Db    680 agcctccaggcctttgtggctgtagccgcccgggtctttgcccatggagcagcaac---- 735

Qy    677 ACTGAGCCCTTAGTGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATG 731
           ||||| || ||  | || ||||| |||||||   | ||| |  ||||||  |  | | |
Db    736 cctgagtccctaaaggcagcagctcaaggatggcactcagatctccatggcccagcaagg 795

Qy    732 ATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
           ||||  |||||| || ||||| || ||   ||  ||       |||||  | |
Db    796 ccaagataaatctaccaccccaggcacctgtgagccaacaggttaattagtccat 850


RESULT   7
Z08865
ID   Z08865 standard; cDNA; 570 BP.
XX
AC   Z08865;
XX
DT   04-NOV-1999  (first entry)
XX
DE   Human interleukin B30 encoding cDNA.
XX
KW   DNAX soluble receptor subunit 1; DNAX cytokine receptor subunit 1;
KW   interleukin B30; DSRS1; DCRS1; IL-B30; cytokine receptor; diagnosis;
KW   inflammatory disorder; inflammatory response; innate immunity;
KW   morphogenic development; immunological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..570
FT                   /*tag=  a
FT                   /product= "IL-B30"
FT   sig_peptide     1..63
FT                   /*tag=  b
FT   mat_peptide     64..567
FT                   /*tag=  c
XX
PN   WO9940195-A1.
XX
PD   12-AUG-1999.
XX
PF   05-FEB-1999;   99WO-US02600.
XX
PR   13-MAY-1998;   98US-0078194.
PR   06-FEB-1998;   98US-0073941.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Kastelein RA,  Mattson JD,  McClanahan TK;
XX
DR   WPI; 1999-527306/44.
DR   P-PSDB; Y29783.
XX
PT   New receptor subunits useful in the treatment inflammatory disorders
XX
PS   Claim 16; Page 26-27; 133pp; English.
XX
CC   The present invention describes a composition (I) comprising DNAX
CC   cytokine receptor subunit I (DCRS1) protein and DNAX soluble receptor
CC   subunit I (DSRS1) protein, which together encode a new mammalian
CC   cytokine-related receptor (R), or DCRS1 and interleukin B30 (IL-B30)
CC   proteins, or DSRS1 and IL-B30 proteins. (I) comprising DSRS1 and DCRS1
CC   is useful for screening for ligands (i.e. agonists/antagonists) from
CC   a library of compounds, which are useful for modulating the physiology
CC   or development of a cell or tissue culture e.g. inflammatory responses,
CC   innate immunity and/or morphogenic development. (R), antibodies and
CC   ligands are useful for treatment of conditions, especially immunological
CC   disorders, associated with conditions exhibiting abnormal expression of
CC   (R). (R) is useful as a phosphate labeling enzyme to label substrates,
CC   and the subunits DSRS1 and DCRS1 are useful as immunogens for generating
CC   antibodies, or as antigens for binding antibodies. Nucleic acids
CC   encoding (R) are useful for identifying related DNAs and mRNAs, and
```

```
CC   variants from other individuals or species. The present sequence
CC   encodes the specifically claimed human IL-B30, for use in the
CC   composition of the present invention.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;
```

```
  Query Match             32.5%;  Score 391;  DB 20;  Length 570;
  Best Local Similarity   81.8%;  Pred. No. 9.1e-109;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172
          |||||||   |||||||  |||||  ||||  |  ||||  ||||||||    ||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232
          |||||||||  ||  |  |||||  ||||  ||||  ||||||||||||||||  ||  |  |||  |||
Db     61 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292
          ||||  ||  |||||||  ||||||  |||  ||  ||||||  |||  ||   ||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggat---ctaagagaagag 177

Qy    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352
          |  ||||||||||||  ||||  ||||  ||||  ||||||||||  ||||||  ||||||||
Db    178 ggagatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgacccc 237

Qy    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412
          ||||||||||  |||||||||  ||||||||||||||||||||  |||  ||||||  ||||
Db    238 caaggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgattttt 297

Qy    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472
          |||  ||  |  |||||  |   ||  ||  ||  ||||  ||||||||||  ||||  ||||||||||||
Db    298 tatgagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagc 357

Qy    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532
          ||  |||  |||  |||||  ||||||||||  ||  ||||||||||||||  |||||  ||||  |
Db    358 cctgtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggt 417

Qy    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592
          ||||  |  |||||||  ||  |||||  ||  |||||  |||  |  ||  ||||  ||||||||  |  |
Db    418 caccactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctc 477

Qy    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652
          ||||||||  |  |||  ||||||||  ||||||||||||||  ||||  ||||  |||||||||
Db    478 cttctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtc 537

Qy    653 TTTGCCCACGGAGCAGCAACTCTGACT 679
          ||||||||  ||||||||||  ||||  |
Db    538 tttgcccatggagcagcaaccctgagt 564
```

```
RESULT   8
X17786
ID   X17786 standard; cDNA; 570 BP.
XX
AC   X17786;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Human interleukin-B30 (IL-B30) polypeptide encoding cDNA.
XX
KW   Cytokine; interleukin-B30; IL-B30; forensic science; cell proliferation;
KW   inflammatory condition; drug screening; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..570
FT                   /*tag=  a
FT                   /product= "IL-B30"
FT   sig_peptide     1..63
FT                   /*tag=  b
FT   mat_peptide     64..567
FT                   /*tag=  c
XX
PN   WO9905280-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;   98WO-US15423.
XX
PR   25-JUL-1997;   97US-0900905.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 1999-142935/12.
DR   P-PSDB; W95002.
XX
PT   Newly isolated or recombinant polynucleotide encoding mammalian
PT   cytokine interleukin-B30 (IL-B30), including fragments - useful for
PT   regulating activation, development, differentiation and function of
PT   various cell types, and for diagnosing and treating conditions
PT   associated with IL-B30
XX
PS   Claim 3; Page 8-9; 83pp; English.
XX
CC   This cDNA encodes a human cytokine interleukin-B30 (IL-B30) polypeptide.
CC   Host cells containing a vector comprising the IL-B30 nucleic acid are
CC   used for the recombinant production of the protein. The polynucleotides
CC   are useful for diagnosis of IL-B30 mediated conditions, and forensic
CC   science (e.g. to distinguish rodent from human, or as a marker to
CC   distinguish between different cells exhibiting differential expression or
CC   modification patterns). The IL-B30 (including fragments), together with
CC   antibodies that bind to IL-B30 are useful for teaching purposes. They
CC   are also used for treating conditions associated with abnormal physiology
CC   or development, including inflammatory conditions. The polypeptide
CC   cytokine should mediate cytokine synthesis and proliferation in cells.
CC   IL-B30 is useful for drug screening to identify compounds having binding
CC   affinity to IL-B30.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;
```

```
  Query Match             32.5%;  Score 391;  DB 20;  Length 570;
  Best Local Similarity   81.8%;  Pred. No. 9.1e-109;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172
          |||||||   |||||||  |||||  ||||  |  ||||  ||||||||    ||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232
          |||||||||  ||  |  |||||  ||||  ||||  ||||||||||||||||  ||  |  |||  |||
Db     61 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292
          ||||  ||  |||||||  ||||||  |||  ||  ||||||  |||  ||   ||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggat---ctaagagaagag 177

Qy    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352
          |  ||||||||||||  ||||  ||||  ||||  ||||||||||  ||||||  ||||||||
Db    178 ggagatgaagagactacaaatgatgttccccatatccagtgtggagatggctgtgacccc 237

Qy    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412
          ||||||||||  |||||||||  ||||||||||||||||||||  |||  ||||||  ||||
Db    238 caaggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgattttt 297

Qy    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472
          |||  ||  |  |||||  |   ||  ||  ||  ||||  ||||||||||  ||||  ||||||||||||
Db    298 tatgagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagc 357

Qy    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532
          ||  |||  |||  |||||  ||||||||||  ||  ||||||||||||||  |||||  ||||  |
```

```
Db    358 cctgtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggt 417

Qy    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592
          |||| | ||||||| || ||||| || ||||| ||| | || ||||  |||||||| | |
Db    418 caccactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctc 477

Qy    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652
          |||||||| | ||| |||||||| ||||||||||||||| ||||  |||| |||||||||
Db    478 cttctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtc 537

Qy    653 TTTGCCCACGGAGCAGCAACTCTGACT 679
          |||||||| |||||||||| |||| |
Db    538 tttgcccatggagcagcaaccctgagt 564


RESULT   9
A52577
ID   A52577 standard; cDNA; 570 BP.
XX
AC   A52577;
XX
DT   27-SEP-2000  (first entry)
XX
DE   Human interleukin-B30 (IL-B30) cDNA.
XX
KW   Interleukin-B30; IL-B30; human; cytokine; cellular signalling;
KW   immune response; haematopoietic cell; IL-6 homologue; GCSF homologue;
KW   granulocyte colony stimulating factor; immune disorder;
KW   inflammatory disease; autoimmune disease; antigen; antibody; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..570
FT                   /*tag= a
FT                   /product= "Human IL-B30 precursor"
FT   sig_peptide     1..63
FT                   /*tag= b
FT   mat_peptide     64..567
FT                   /*tag= c
FT                   /product= "Mature human IL-B30"
XX
PN   US6060284-A.
XX
PD   09-MAY-2000.
XX
PF   24-JUL-1998;   98US-0122443.
XX
PR   25-JUL-1997;   97US-0053765.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF;
XX
DR   WPI; 2000-364420/31.
DR   P-PSDB; B01981.
XX
PT   Novel recombinant DNA encoding cytokines especially interleukin-B30
PT   useful as probes or primers for diagnosing immune disorders including
PT   autoimmune or chronic inflammatory conditions  -
XX
PS   Claim 3; Column 5-6; 32pp; English.
XX
CC   This sequence represents cDNA encoding human interleukin-B30 (IL-B30).
CC   IL-B30 is a novel cytokine, exhibiting significant homology to IL-6 and
CC   GCSF (granulocyte colony stimulating factor). Cytokines play a critical
CC   role in signalling between immune or other cells during an immune
CC   response. The precise role if IL-B30 is not yet known - it is likely to
CC   have either a stimulatory or an inhibitory effect on haematopoietic cells
CC   such as T-cells, B-cells, natural killer (NK) cells and macrophages.
CC   Alternatively, it may affect vascular physiology or development, or have
CC   neuronal effects. IL-B30, its fragments, IL-B30 nucleotides, agonists and
CC   antagonists are useful in the diagnosis and treatment of disorders
CC   associated with abnormal expression or activity of IL-B30 e.g.,
CC   inflammatory or autoimmune disorders. Nucleotides encoding IL-B30 are
CC   useful for recombinant expression of IL-B30 in a host cell, and as a
CC   source of probes and primers. The IL-B30 probes and primers can be used
CC   to detect levels of IL-B30 expression in samples from patients suspected
CC   of having an inflammatory or autoimmune disorder. IL-B30 nucleotides may
CC   also be used to identify homologous genes in other species. IL-B30
CC   protein or its fragments are useful as antigens for raising antibodies to
CC   various linear and conformational epitopes. Such antibodies may be used
CC   to detect levels of IL-B30 protein in a sample.
XX
SQ   Sequence 570 BP; 121 A; 168 C; 157 G; 124 T; 0 other;


  Query Match             32.5%;  Score 391;  DB 21;  Length 570;
  Best Local Similarity   81.8%;  Pred. No. 9.1e-109;
  Matches  464;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy    113 ATGCTGGATTGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTG 172
          |||||||   |||||| |||||  |||| | |||| ||||||||    ||||||||
Db      1 atgctggggagcagagctgtaatgctgctgttgctgctgccctggacagctcagggcaga 60

Qy    173 GCTGTGCCTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTC 232
          ||||||||| || | ||||| |||| |||| ||||||||||||||||| || | ||| |||
Db     61 gctgtgcctgggggcagcagccctgcctggactcagtgccagcagctttcacagaagctc 120

Qy    233 TGCATGCTAGCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAA 292
          ||||  || |||||||  |||||| ||| || |||||| ||| ||   ||||||||||
Db    121 tgcacactggcctggagtgcacatccactagtgggacacatggat---ctaagagaagag 177

Qy    293 GAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCA 352
          |  |||||||||||||| |||| |||| |||| |||||||||| |||||| |||||||
Db    178 ggagatgaagagactacaaatgatgttcccatatccagtgtggagatggctgtgacccc 237

Qy    353 CAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTT 412
          |||||||||| |||||||| |||||||||||||||||||||| ||| |||||| ||||
Db    238 caaggactcagggacaacagtcagttctgcttgcaaaggatccaccagggtctgattttt 297

Qy    413 TATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTACTCCCTGATAGC 472
          ||| || | ||||| |  || || || |||| |||||||||| |||| ||||||||||||
Db    298 tatgagaagctgctaggatcggatattttcacaggggagccttctctgctccctgatagc 357

Qy    473 CCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTCCAGCCAGAGGAT 532
          ||  ||| ||| ||||| |||||||||| || |||||||||||||| ||||| |||| |
Db    358 cctgtggcgcagcttcatgcctccctactgggcctcagccaactcctgcagcctgagggt 417

Qy    533 CACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCC 592
          |||| | ||||||| || ||||| || ||||| ||| | || ||||  |||||||| | |
Db    418 caccactgggagactcagcagattccaagcctcagtcccagccagccatggcagcgtctc 477

Qy    593 CTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTC 652
          |||||||| | ||| |||||||| |||||||||||||| ||||  |||| |||||||||
Db    478 cttctccgcttcaaaatccttcgcagcctccaggcctttgtggctgtagccgcccgggtc 537

Qy    653 TTTGCCCACGGAGCAGCAACTCTGACT 679
          |||||||| |||||||||| |||| |
Db    538 tttgcccatggagcagcaaccctgagt 564


RESULT  10
A42676
ID   A42676 standard; cDNA; 412 BP.
XX
AC   A42676;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1416.
XX
KW   Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
```

```
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021990-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24205.
XX
PR   15-OCT-1998;   98US-0104435.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI   Merberg D, Treacy M;
XX
DR   WPI; 2000-317937/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 459; 618pp; English.
XX
CC   A41261 to A43419 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include: chemotactic;
CC   proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;
CC   haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;
CC   antifungal; antiviral; antidiabetic; antiasthmatic; vulnerary; antiulcer;
CC   osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;
CC   cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be
CC   used for gene therapy and in vaccines. The sESTs are useful as probes for
CC   the identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression
CC   and psoriasis. A43420 to A43425 represent linker variants which are
CC   given in the exemplification of the present invention.
XX
SQ   Sequence 412 BP; 115 A; 101 C; 130 G; 66 T; 0 other;

  Query Match             16.9%;  Score 203.2;  DB 21;  Length 412;
  Best Local Similarity   77.9%;  Pred. No. 8.6e-52;
  Matches  271;  Conservative    0;  Mismatches   73;  Indels    4;  Gaps    2;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      68 agagaatcaggctcaaagcaagtggaagtgggcagagattccaccaggactggtgcaagg 127

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || ||||||||||   ||||||| |||||
Db     128 cgcagagccagccagatttgagaagaaggcaaaaagatgctggggagcagagctgtaatg 187

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
           |||| | |||| ||||||||    ||||||||  ||||||||| || | ||||| |||
Db     188 ctgctgttgctgctgccctggacagctcagggcagagctgtgcctgggggcagcagccct 247

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| |||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     248 gcctggactcagtgccagcagctttcacagaagctctgcacactggcctggagtgcacat 307

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
           ||| || |||||| ||| ||   |||||||||| | ||||||||||||| |||| ||
Db     308 ccactagtgggacacatggat---ctaagagaagagggagatgaagagactacaaatgat 364

Qy     317 GTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCAAG 364
           || |||| |||||||||| |||||| |||||||| | || |||| ||
Db     365 gttccccatatccagtgtggagatggctgtgacccccagaactcgag 412


RESULT  11
Z37263
ID   Z37263 standard; DNA; 2398 BP.
XX
AC   Z37263;
XX
DT   01-FEB-2000  (first entry)
XX
DE   SGRF coding sequence.
XX
KW   SGRF; human; Interleukin-6 G-CSF related factor; cell proliferation;
KW   immune system; haematopoietic system; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9954357-A1.
XX
PD   28-OCT-1999.
XX
PF   14-APR-1999;   99WO-JP01997.
XX
PR   14-APR-1998;   98JP-0121805.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Hirata Y;
XX
DR   WPI; 2000-013230/01.
DR   P-PSDB; Y54606.
XX
PT   Novel cytokine-like protein, with activity of supporting proliferation
PT   of myeloid cells, useful in treating abnormality of cell proliferation
PT   in immune and haematopoiesis systems  -
XX
PS   Disclosure; Fig 6; 69pp; Japanese.
XX
CC   This sequence encodes the Interleukin-6 G-CSF related factor (SGRF)
CC   protein of the invention. The protein is a member of the IL-6/G-CSF/MSF
CC   family. The protein can be used in drugs for treating diseases due to
CC   abnormality of cell proliferation in the immune system and haematopoietic
CC   system.
XX
SQ   Sequence 2398 BP; 599 A; 560 C; 633 G; 606 T; 0 other;

  Query Match             12.6%;  Score 151.6;  DB 21;  Length 2398;
  Best Local Similarity   76.7%;  Pred. No. 1.1e-35;
  Matches  198;  Conservative    0;  Mismatches   59;  Indels    1;  Gaps    1;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db     365 agagaatcaggctcaaagcaagtggaagtgggcagagattccaccaggactggtgcaagg 424

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || ||||||||||   ||||||| |||||
Db     425 cgcagagccagccagatttgagaagaaggcaaaaagatgctggggagcagagctgtaatg 484
```

```
Qy    137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
          |||| | |||| ||||||||    ||||||||   |||||||| || | ||||| |||
Db    485 ctgctgttgctgctgccctggacagctcagggcagagctgtgcctgggggcagcagccct 544

Qy    197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
          | |||| ||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db    545 gcctggactcagtgccagcagctttcacagaagctctgcacactggcctggagtgcacat 604

Qy    257 GCACCAGCGGGACATATG 274
          ||| || |||||| |||
Db    605 ccactagtgggacacatg 622


RESULT 12
Z43768/c
ID   Z43768 standard; cDNA; 9312 BP.
XX
AC   Z43768;
XX
DT   08-MAR-2000  (first entry)
XX
DE   Human CRP-ductin cDNA.
XX
KW   CRP-ductin; human; SCUZ protein;;cytostatic; treatment; carcinoma;
KW   brain tumor; breast tumor; lung; melanoma; pancreas; endometrium;
KW   oesophagus; colon; stomach; ss.
XX
OS   Homo sapiens.
XX
PN   DE19829660-C1.
XX
PD   23-DEC-1999.
XX
PF   02-JUL-1998;   98DE-1029660.
XX
PR   02-JUL-1998;   98DE-1029660.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI   Mollenhauer J,  Poustka A,  Krebs I;
XX
DR   WPI; 2000-054563/05.
XX
PT   Transgenic non-human mammal useful for testing compounds for anticancer
PT   activity  -
XX
PS   Claim 14; Fig 1; 14pp; German.
XX
CC   This invention describes a novel non-human mammal, comprises a gene,
CC   that codes for a SCUZ protein, that is altered so that it is switched
CC   off, or can be switched off in a time- or tissue-specific manner. The
CC   protein described in the invention has cytostatic activity. The non-human
CC   mammal or cells derived from it are useful for testing of compounds that
CC   influence the function, in particular the activity of a SCUZ protein. The
CC   compounds are useful for treatment of carcinoma, in particular brain or
CC   breast tumors, lung carcinoma, melanoma or pancreatic, endometrial,
CC   oesophageal, colon or stomach carcinomas. This sequence represents a
CC   nucleic acid fragment which encodes the human CRP-ductin gene exon 1 and
CC   its promoter which is used in the method of the invention.
XX
SQ   Sequence 9312 BP; 2266 A; 2240 C; 2140 G; 2664 T; 2 other;


 Query Match             2.9%;  Score 35.4;  DB 21;  Length 9312;
  Best Local Similarity   54.1%;  Pred. No. 4.4;
  Matches   72;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    838 AAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGA 897
          | ||||| ||||||  ||       |  |  ||  ||  |    ||||||  |||| |||
Db    358 AGAGAGAGAAGAAAAGAAGGAGGAGGAGGAGAGAAGGAAAAGAGAAGGAGAAGGAGGAGA 299

Qy    898 AGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGG 957
          |||| | |   .  ||     ||    ||| | ||   ||  |  | | ||   ||||||
Db    298 AGAAGAGAAGAAGAAGAAGGAGGAGGAGGAGGAGGAGGAGGGGAGGGGAGGGGGGAGGGG 239

Qy    958 AAGGGGGCGGTGG 970
          |||||| |  ||
Db    238 GAGGGGGAGAGGG 226


RESULT 13
C06417
ID   C06417 standard; cDNA; 396 BP.
XX
AC   C06417;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 10492.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 10492; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 396 BP; 129 A; 82 C; 79 G; 104 T; 2 other;


  Query Match             2.9%;  Score 35;  DB 21;  Length 396;
  Best Local Similarity   54.2%;  Pred. No. 1.1;
  Matches   71;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy    671 ACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTCCCATGGCTACCAT 730
          ||   | ||  ||  |||||||| | ||  ||   | |   ||    |  |   ||| |
Db     34 acagcgtctaggcttttagtgcctctgccaaatcgttcttcgggaagccagtkcaccttt 93

Qy    731 GATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTTGTGC 790
            || | |||| ||| ||   | |||| |||| ||   || ||  ||      || ||
Db     94 tccaaaaataatgtatgagagaatacatttacctgtaccttcactaatrctagattctgt 153

Qy    791 TGTTCTTGTTT 801
```

```
                 || | |||| |
Db     154 tgcttttgtgt 164


RESULT  14
G09932
ID   C09932 standard; cDNA; 292 BP.
XX
AC   C09932;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14007.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 14007; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 292 BP; 75 A; 98 C; 49 G; 67 T; 3 other;

  Query Match           2.9%;  Score 34.8;  DB 21;  Length 292;
  Best Local Similarity   50.0%;  Pred. No. 1;
  Matches   84;  Conservative    1;  Mismatches   83;  Indels    0;  Gaps    0;

Qy     498 TACTAGGACTCAGCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGC 557
           ||||    | |||| |    |  ||||   || |  | | :| | ||   |||||   ||
Db       3 tactgattcacagcgagaggcagcagcagcagtagcascagcagcagcaacaacagc 62

Qy     558 CCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAA 617
           | || ||      |   || | ||  |    | | |  |    |||  |   ||||  |
Db      63 gctgcttgtcacgaatcgaggattgcaatgagctcatcgttttctccttgcagcttcaca 122

Qy     618 GCCTCCAGGCCTTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAG 665
           |  ||  | |||  |||| |  |||||    |||||  | |||  ||
Db     123 actgccctggcttcctggctactgctgctgctgtcttctcacacatag 170


RESULT  15
X83478
ID   X83478 standard; cDNA to mRNA; 3960 BP.
XX
AC   X83478;
XX
DT   14-SEP-1999  (first entry)
XX
DE   A.thaliana potassium channel KAT2 gene.
XX
KW   Alteration; activity; potassium transporter; AtKT family; KAT1; AKT1;
KW   inwardly rectifying potassium channel; AtKC1; transgenic plant; uptake;
KW   apical dominance; overexpression; characteristic; agricultural trait; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9931259-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-EP08190.
XX
PR   15-DEC-1997;   97EP-0122104.
XX
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI   Palme K,  Reintanz B,  Saedler H,  Wisman E;
XX
DR   WPI; 1999-395189/33.
XX
PT   Transgenic plants having altered potassium levels
XX
PS   Disclosure; Fig 5; 41pp; English.
XX
CC   The invention relates to an alteration in the activity of at least one
CC   plant potassium transporter of the AtKT family and/or at least one type
CC   of plant inwardly rectifying potassium channel (Kin+) of the KAT1, AKT1
CC   or the AtKC1 family in transgenic plants which leads to a reduced apical
CC   dominance, and reduced K+ uptake. Overexpression of one of these proteins
CC   leads to improved plant characteristic.  The sequence presented here
CC   corresponds to the Arabidopsis thaliana KAT2 gene sequence.
CC   The method can be used to produce transgenic plants that have reduced
CC   or increased uptake of potassium, depending on the agricultural trait
CC   needed.
XX
SQ   Sequence 3960 BP; 1264 A; 772 C; 682 G; 1242 T; 0 other;

  Query Match           2.9%;  Score 34.6;  DB 20;  Length 3960;
  Best Local Similarity   55.4%;  Pred. No. 4.8;
  Matches   67;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Qy     766 TTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAG 825
           || ||  |     | |  ||  ||| | ||||   || ||||  || | |  |  | |||
Db    3399 ttgatggaaaaggttgctctgatgcagatcttacctcatttgaatttcattcacaggaag 3458

Qy     826 GACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGG 885
           | | |  ||||  |||  ||| |  |||| |  | | || | |||||   |   ||| |
Db    3459 catatccatattgcagatcaaacatccaaatcaaacaacatgaagctgcaaaaccaaaag 3518

Qy     886 A 886
           |
Db    3519 a 3519


Search completed: June 21, 2001, 02:09:22
Job time: 4611 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2001, 01:24:55 ; Search time 1723.79 Seconds
(without alignments)
6596.965 Million cell updates/sec

Title: US-09-558-474-3
Perfect score: 1203
Sequence: 1 CGCTTAGAAGTCGGACTACA.....TACCAATACGAACTGACAAA 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	638	53.0	1150	137	BE572471	BE572471 601331259
2	481.4	40.0	484	141	BE850725	BE850725 ux01d01.y
3	479	39.8	897	137	BE571481	BE571481 601333907
4	429	35.7	430	141	BE847383	BE847383 uw26b04.y
5	384	31.9	484	166	BE335675	BE335675 us87h10.y
c 6	368	30.6	523	147	BF318114	BF318114 ux01d01.x
7	340	28.3	469	141	BE851424	BE851424 uw95h11.y
8	299.8	24.9	833	104	AJ277049	AJ277049 AJ277049
9	298	24.8	323	129	BB270313	BB270313 BB270313
10	295.2	24.5	320	130	BB274745	BB274745 BB274745
c 11	295	24.5	458	143	BF018105	BF018105 uw95h11.x
12	290.8	24.2	492	7	AA418955	AA418955 zw01c10.r
c 13	281	23.4	447	123	AW987117	AW987117 uf14d07.x
14	183.8	15.3	372	155	C06368	C06368 C06368 Huma
c 15	172.4	14.3	542	102	AI796983	AI796983 we25e07.x
16	151.8	12.6	719	153	BG435629	BG435629 602506911
c 17	135.4	11.3	496	175	BG236314	BG236314 naf26b02.
c 18	121.6	10.1	575	31	AV598349	AV598349 AV598349
c 19	120	10.0	367	145	BF189762	BF189762 235459 MA
c 20	109.6	9.1	399	114	AW316433	AW316433 14841 MAR
c 21	108.8	9.0	457	117	AW511070	AW511070 hd38e10.x
c 22	96.8	8.0	566	165	BE237060	BE237060 146167 MA
23	78.8	6.6	137	188	T18577	T18577 hbc2125 Hum
c 24	71.2	5.9	400	14	AI004049	AI004049 ou11d12.x
c 25	46	3.8	1101	219	CNS01784	AL107710 Drosophil
c 26	45.8	3.8	997	219	CNS005TE	AL060767 Drosophil
c 27	45.6	3.8	369	111	AW072323	AW072323 wz99h07.x
c 28	45.4	3.8	943	219	CNS00ZIM	AL097720 Drosophil
c 29	44.8	3.7	987	219	CNS00418	AL066537 Drosophil
30	44.8	3.7	990	219	CNS017DC	AL107898 Drosophil
c 31	43.4	3.6	842	256	B19240	B19240 F9C2-T7 IGF
c 32	42.8	3.6	397	7	AA418747	AA418747 zw01c10.s
c 33	42.6	3.5	1101	219	CNS017GS	AL108022 Drosophil
34	42.4	3.5	1101	219	CNS0181N	AL108773 Drosophil
35	42	3.5	1101	219	CNS017RP	AL108415 Drosophil
c 36	41.8	3.5	698	32	AV728560	AV728560 AV728560
37	40.6	3.4	305	229	AQ512257	AQ512257 HS_5081_A
38	40.4	3.4	957	105	AL526604	AL526604 AL526604
39	40.2	3.3	937	142	BE964027	BE964027 601657796
c 40	40.2	3.3	1100	219	CNS016KD	AL106855 Drosophil
41	40	3.3	2092	192	AK018120	AK018120 Mus muscu
c 42	39.8	3.3	799	80	BF262498	Bf262498 HV_CEa000
c 43	39.8	3.3	1159	219	CNS015XR	AL106041 Drosophil
c 44	39.6	3.3	633	256	B19596	B19596 T30K12-T7.1
45	39.6	3.3	1101	219	CNS006BP	AL064052 Drosophil

ALIGNMENTS

RESULT 1
BE572471
LOCUS BE572471 1150 bp mRNA EST 15-AUG-2000
DEFINITION 601331259F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3708687 5', mRNA sequence.
ACCESSION BE572471
VERSION BE572471.1 GI:9816191
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1150)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8928 row: f column: 16
High quality sequence stop: 716.
FEATURES Location/Qualifiers
source 1. .1150
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3708687"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 330 a 259 c 329 g 231 t 1 others
ORIGIN

Query Match 53.0%; Score 638; DB 137; Length 1150;
Best Local Similarity 89.7%; Pred. No. 1.8e-174;
Matches 788; Conservative 0; Mismatches 75; Indels 15; Gaps 9;

```
Qy      37 CAAAGGAGGTGGATAGGGGGTCCACAGGCCTGGTGCAGATC-ACAGAGCCAGCCAGATCT 95
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1 CAAAGGAGGTGGATAGGGGGTCCACAGGCCTGGTGCAGATCGACAGAGCCAGCCAGATCT 60

Qy      96 GAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATAATGCTA-TGGCTGTTGCCC 154
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      61 GAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATAATGCTAGTGGCTGTTGCCC 120

Qy     155 TGGGTCACTCAGGGCC-TGGCTGTGCC-TAGGAGTAGCAGTCCTGACTGGGCTCAGTGCC 212
           |||||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db     121 TGGGTCACTCAGGGCCGTGGCTGTGCCGTAGGAGTAGCAGTCCTGACTGGGCTCAGTGCC 180

Qy     213 AGCAGCTCTCTCGGAATCTCTGC--ATGCTAGCCTGGAACGCACATGCACCAGCGGGACA 270
           |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     181 AGCAGCTCTCTCGGAATCTCTGCGATGCGTAGCCTGGAACGCACATGCACCAGCGGGACA 240

Qy     271 TATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAATGTGCCCCCGTATCCA 330
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAATGTGCCCCCGTATCCA 300

Qy     331 GTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAG 390
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GTGTGAAGATGGTTGTGACCCACAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAG 360

Qy     391 GATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGA 450
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GATCCGCCAAGGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGA 420

Qy     451 GCCTGCTCTACTCCC-TGATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCA 509
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     421 GCCTGCTCTACTCCCGTGATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCA 480

Qy     510 GCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCC-TGAGT 568
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     481 GCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTTGAGT 540

Qy     569 TCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCC 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCC 600
```

```
Qy    629 TTTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTA 688
          ||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||||
Db    601 TTTTTGGCCATAGCTGCCCGGGTCTTGGGCCACGGAGCAGCAACTCTGACTGAGCCCTTA 660

Qy    689 GTGCCAACAGCTTAAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCA 748
          |||| |     | | ||||||| |||||||||||| ||||||||||||  |
Db    661 GTGCAACAGCTTAAGGGATGCCAAGGTTCCCATGGGTACCATGATAAGATAATCTATCAG 720

Qy    749 GCCCAGACATCTACCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGT 808
          ||| ||||||||||||||||| | |||  |  |    || |    |      | |   |
Db    721 GCCAAGACATCTACCAGTTAAGTTACCATTAGGATTGTGGGGAACTGGAAGGTGGAGCG- 779

Qy    809 TTTGCGTGAAGGGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCA-GG 867
          ||||||||||| ||   ||||||| |||||||| | ||||| |||  ||||| ||
Db    780 -----GTGAAGGGCAGGGGACCCATTATCAAAGAGAAAGGTAACAACCCCAGGAGCAGGG 834

Qy    868 CAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGAATAAA 905
          |||| ||  | | |||||||   | |  ||||| |||
Db    835 CAGCGGGTAACACAAAGGAGACTGTGTCAAAGAAGAAA 872


RESULT  2
BE850725
LOCUS       BE850725      484 bp    mRNA            EST       26-SEP-2000
DEFINITION  ux01d01.y1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3470209 5', mRNA sequence.
ACCESSION   BE850725
VERSION     BE850725.1  GI:10309084
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1389569
            Seq primer: -40RP from Gibco
            High quality sequence stop: 459.
FEATURES             Location/Qualifiers
     source          1. .484
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3470209"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      144 a    109 c    130 g    100 t      1 others
ORIGIN


  Query Match           40.0%;  Score 481.4;  DB 141;  Length 484;
  Best Local Similarity  99.6%;  Pred. No. 5.3e-129;
  Matches  482;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy    658 CCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTC 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTTAAGGATGCCCAGGTTC 60

Qy    718 CCATGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCA 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCATGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTACCAGTTAATTAACCCA 120

Qy    778 TTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATT 837
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    121 TTAGGACTTGTGCTGTTCTTGTTTTGTTTGTTTTGCGTGAAGGGCAAGGACACCATTATT 180

Qy    838 AAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGA 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGA 240

Qy    898 AGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGG 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGG 300

Qy    958 AAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCC 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCC 360

Qy   1018 AGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGT 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGT 420

Qy   1078 GGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTC 1137
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGCTCANGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTC 480

Qy   1138 TGTC 1141
          ||||
Db    481 TGTC 484


RESULT  3
BE571481
LOCUS       BE571481      897 bp    mRNA            EST       15-AUG-2000
DEFINITION  601333907F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3711529 5',
            mRNA sequence.
ACCESSION   BE571481
VERSION     BE571481.1  GI:9815201
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8935  row: m  column: 02
            High quality sequence stop: 597.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3711529"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
```

```
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      239 a    201 c    262 g    194 t      1 others
ORIGIN


  Query Match             39.8%;  Score 479;  DB 137;  Length 897;
  Best Local Similarity   93.5%;  Pred. No. 3.2e-128;
  Matches  609;  Conservative    0;  Mismatches   30;  Indels   12;  Gaps   10;

Qy       2 GCTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCAC 61
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 GGTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCAC 61

Qy      62 AGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGAT 121
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      62 AGGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCA-GGAACAAGATGCTGGA- 119

Qy     122 TGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCT 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     120 TGCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGG-TGTGCCT 178

Qy     182 AGGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTA 241
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     179 AGGAGTAGCAGTCCTGAGTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTA 238

Qy     242 GCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAA 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     239 GCCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAA 298

Qy     302 GAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTC 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 GAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTC 358

Qy     362 AAGGACAACAGCCAGTTCTGCTTGC-AAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 420
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     359 AAGGACAACAGCCAGTTCTGCTTGCAAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCA 418

Qy     421 CCTGCTTGACTCTGACATCTTCAA-AGGGGAGCCTGCTCTACTCCCTGATAGCCCCATGG 479
          |||||||||||||||||||||||| ||||||||   ||||||||||||   |||   |
Db     419 CCTGCTTGACTCTGACATCTTCAAGAGGGGAGCCGTGTCTACTCCCTGATTAGCCCCATG 478

Qy     480 AGCAACTTCACACCTCCCTACTAGGACTCAGCC-AACTCCTCCAGCCAGAGGATCACCCC 538
             | |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     479 GAGCAATTCACACCTCCCTACTAGGACTCAGCCAAACTCCTCCAGCCAGAGGATCACCCC 538

Qy     539 CGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGTGGCAGCGCCCCCTTCTC 598
          | |||||||||||||||||| ||||||||||||||||||||      |||||||||
Db     539 C-GGAGACCCAACAGATGCCCAG-CTGAGTTCTAGTCAGCAGTGG--AGCGCCCCTTCTC 594

Qy     599 CGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAGCTGCCCGG 649
          ||||||||||| |   | ||||||| |||| ||  |||||||||| ||
Db     595 CGTTCCAAGATC--TTCGAAGCTCCAGGGCTTTGTGCGCATAGCTGCCGGG 643


RESULT   4
BE847383
LOCUS       BE847383      430 bp    mRNA            EST       26-SEP-2000
DEFINITION  uw26b04.y1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3417775 5', mRNA sequence.
ACCESSION   BE847383
VERSION     BE847383.1  GI:10305722
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1093587
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3417775"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      120 a     99 c    123 g     88 t
ORIGIN


  Query Match             35.7%;  Score 429;  DB 141;  Length 430;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-114;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 CTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCACA 62
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTTAGAAGTCGGACTACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCACA 60

Qy      63 GGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATT 122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGCCTGGTGCAGATCACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATT 120

Qy     123 GCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTA 182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCAGAGCAGTAATAATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTA 180

Qy     183 GGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAG 242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGAGTAGCAGTCCTGACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAG 240

Qy     243 CCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAG 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTGGAACGCACATGCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAG 300

Qy     303 AGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCA 362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGATGGTTGTGACCCACAAGGACTCA 360

Qy     363 AGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACC 422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAAGGTCTGGCTTTTTATAAGCACC 420

Qy     423 TGCTTGACT 431
          |||||||||
Db     421 TGCTTGACT 429


RESULT   5
```

```
BE335675
LOCUS       BE335675      484 bp    mRNA            EST       14-JUL-2000
DEFINITION  us87h10.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:3325411 5', mRNA sequence.
ACCESSION   BE335675
VERSION     BE335675.1  GI:9209451
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1069575
            Seq primer: -40RP from Gibco
            High quality sequence stop: 460.
FEATURES             Location/Qualifiers
     source          1. .484
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3325411"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      161 a     97 c    124 g    102 t
ORIGIN


  Query Match           31.9%;  Score 384;  DB 166;  Length 484;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-100;
  Matches  384;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     820 GGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGA 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCAAGGACACCATTATTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGA 60

Qy     880 GAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCA 939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCA 120

Qy     940 GCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGC 999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGC 180

Qy    1000 ATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAA 1059
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAA 240

Qy    1060 CATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAG 1119
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAG 300

Qy    1120 CTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACA 1179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACA 360

Qy    1180 GGACTACCAATACGAACTGACAAA 1203
           |||||||||||||||||||||||
Db     361 GGACTACCAATACGAACTGACAAA 384


RESULT   6
BF318114/c
LOCUS       BF318114      523 bp    mRNA            EST       21-NOV-2000
DEFINITION  ux01d01.x1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3470209 3', mRNA sequence.
ACCESSION   BF318114
VERSION     BF318114.1  GI:11266662
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: ux01d01.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1389569
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3470209"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      130 a    127 c     96 g    170 t
ORIGIN


  Query Match           30.6%;  Score 368;  DB 147;  Length 523;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-96;
  Matches  368;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     836 TTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     523 TTAAAGAGAAAAGAAACAAACCCCAGAGCAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAA 464

Qy     896 GAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGG 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     463 GAAGAATAAAGTCTCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGG 404

Qy     956 GGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGC 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     403 GGAAGGGGGCGGTGGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGC 344

Qy    1016 CCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     343 CCAGGCTTTGGCCATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGG 284
```

```
Qy   1076 GTGGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGC 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 GTGGCTCAGGGAAATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGC 224

Qy   1136 TCTGTCCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAA 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 TCTGTCCCTCAGTTCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAA 164

Qy   1196 CTGACAAA 1203
          ||||||||
Db    163 CTGACAAA 156
```

```
RESULT   7
BE851424
LOCUS       BE851424      469 bp    mRNA            EST       26-SEP-2000
DEFINITION  uw95h11.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:3469797 5', mRNA sequence.
ACCESSION   BE851424
VERSION     BE851424.1  GI:10309763
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1389157
            Seq primer: -40RP from Gibco
            High quality sequence stop: 396.
FEATURES             Location/Qualifiers
     source          1. .469
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3469797"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      147 a    89 c    118 g    115 t
ORIGIN
```

```
  Query Match             28.3%;  Score 340;  DB 141;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-88;
  Matches  340;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    864 CAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCT 923
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAGGCAGCTGGCTAGAGAAAGGAGCTGGAGAAGAAGAATAAAGTCTCGAGCCCTTGGCCT 60

Qy    924 TGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTG 983
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACTG 120

Qy    984 TGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAA 1043
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAAA 180
```

```
Qy   1044 AACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCACA 1103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCACA 240

Qy   1104 GTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTAG 1163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTAG 300

Qy   1164 TCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 1203
          ||||||||||||||||||||||||||||||||||||||||
Db    301 TCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 340
```

```
RESULT   8
AJ277049
LOCUS       AJ277049      833 bp    mRNA            EST       30-MAR-2000
DEFINITION  AJ277049 Placenta Matchmaker cDNA library(Clontech) Homo sapiens
            cDNA clone BHA2, mRNA sequence.
ACCESSION   AJ277049
VERSION     AJ277049.1  GI:7342435
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   Bach,S., Portetelle,D. and Vandenbol,M.
  TITLE     Isolation of new human EST sequences by using the MSB4 yeast gene
            as bait in the two-hybrid screening of a Human Placenta cDNA
            library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Bach SL
            Microbiology Unit
            Gembloux Agricultural University
            Avenue Marechal Juin 6, B-5030 Gembloux, Belgium.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BHA2"
                     /clone_lib="Placenta Matchmaker cDNA library(Clontech)"
                     /sex="Female"
                     /tissue_type="Placenta"
BASE COUNT      227 a    197 c    200 g    209 t
ORIGIN
```

```
  Query Match             24.9%;  Score 299.8;  DB 104;  Length 833;
  Best Local Similarity   78.3%;  Pred. No. 4.8e-76;
  Matches  400;  Conservative    0;  Mismatches  102;  Indels    9;  Gaps    3;

Qy    281 CTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGAT 340
          |||||||||||| |  |||||||||||||| |||| |||| |||| |||||||||||| ||||
Db    100 CTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGAT 159

Qy    341 GGTTGTGACCCACAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAA 400
          ||  |||||||| |||||||||| |||||||| |||||||||||||||||||||| |||
Db    160 GGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAG 219

Qy    401 GGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTA 460
          ||||||  ||||||| || | ||||| |  || || || |||| |||||||||| ||||
Db    220 GGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTG 279

Qy    461 CTCCCTGATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTC 520
          ||||||||||||||  |||  || |||||  ||||||||||| || ||||||||||||||
Db    280 CTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTG 339

Qy    521 CAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAG 580
          ||||| |||| ||||| | ||||||| || ||||| || ||||| ||| | || ||||
Db    340 CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCA 399

Qy    581 TGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATA 640
```

```
                ||||||||| | ||||||||| | ||| |||||||| |||||||||||||| ||||  ||
Db     400 TGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTA 459

Qy     641 GCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCT 700
           || ||||||||||||||| ||||||||||     ||||| || ||  | || |||||
Db     460 GCCGCCCGGGTCTTTGCCCATGGAGCAGCAAC----CCTGAGTCCCTAAAGGCAGCAGCT 515

Qy     701 TAAGGATG---CCCAGGTTCCCATGG--CTACCATGATAAGACTAATCTATCAGCCCAGA 755
            |||||||   | ||| |  ||||||  |  | | |  |||  |||||| || |||||
Db     516 CAAGGATGGCACTCAGATCTCCATGGCCCAGCAAGGCCAAGATAAATCTACCACCCCAGG 575

Qy     756 CATCTACCAGTTAATTAACCCATTAGGACTT 786
           || ||   ||  ||       |||||  | |
Db     576 CACCTGTGAGCCAACAGGTTAATTAGTCCAT 606


RESULT  9
BB270313
LOCUS       BB270313      323 bp    mRNA            EST       07-JUL-2000
DEFINITION  BB270313 RIKEN full-length enriched, 10 days neonate cortex Mus
            musculus cDNA clone A830047G15 3', mRNA sequence.
ACCESSION   BB270313
VERSION     BB270313.1  GI:8966785
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .323
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A830047G15"
                     /clone_lib="RIKEN full-length enriched, 10 days neonate
                     cortex"
                     /tissue_type="cortex"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT       73 a     99 c     66 g     85 t
ORIGIN


  Query Match             24.8%;  Score 298;  DB 129;  Length 323;
  Best Local Similarity   95.3%;  Pred. No. 1.2e-75;
  Matches  307;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     516 TCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTC 575
           |||  |||||||||| |||||||||||||||||||||||||||||| ||||| || ||||
Db       2 TCCCACAGCCAGAGGTTCACCCCCGGGAGACCCAACAGATGCCCAACCTGAATTTTAGTT 61

Qy     576 AGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGG 635
           |||| ||||||||||||||| ||| ||||||| |||||||||||||||| |||||||||
Db      62 AGCAATGGCAGCGCCCCCTTTTCCCTTCCAAGTTCCTTCGAAGCCTCCAGCCCTTTTTGG 121

Qy     636 CCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAA 695
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||
Db     122 CCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGACCCCTTAGTCCCAA 181

Qy     696 CAGCTTAAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCCAGA 755
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 CAGCTTAAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCCAGA 241

Qy     756 CATCTACCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGT 815
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     242 CATCTACCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTTGTTTGTTTTGCGT 301

Qy     816 GAAGGGCAAGGACACCATTATT 837
           |||||||||||||||||||||
Db     302 GAAGGGCAAGGACACCATTATT 323


RESULT  10
BB274745
LOCUS       BB274745      320 bp    mRNA            EST       07-JUL-2000
DEFINITION  BB274745 RIKEN full-length enriched, 10 days neonate cortex Mus
            musculus cDNA clone A830089H09 3', mRNA sequence.
ACCESSION   BB274745
VERSION     BB274745.1  GI:8971766
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
```

```
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .320
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A830089H09"
                     /clone_lib="RIKEN full-length enriched, 10 days neonate
                     cortex"
                     /tissue_type="cortex"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT       71 a    101 c     63 g     85 t
ORIGIN
```

```
  Query Match             24.5%;  Score 295.2;  DB 130;  Length 320;
  Best Local Similarity   95.9%;  Pred. No. 7.5e-75;
  Matches  303;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy     522 AGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAGT 581
           | |||| |||||||||| ||||||||||||||| |||| ||||||||| |||| ||||||
Db       5 ATCCAGTGGATCACCCCCCGGAGACCCAACAGATTCCCACCCTGAGTTTTAGTTAGCAGT 64

Qy     582 GGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATAG 641
           |||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db      65 GGCAGCGCCCCCTTCTCCCTTCCAAGATCCTTCGAAGCCTCCAGCCCTTTTTGGCCATAC 124

Qy     642 CTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTGCCAACAGCTT 701
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     125 CTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAGTCCCAACAGCTT 184

Qy     702 AAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTA 761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 AAGGATGCCCAGGTTCCCATGGCTACCATGATAAGACTAATCTATCAGCCCAGACATCTA 244

Qy     762 CCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTCGTTTGTTTTGCGTGAAGGG 821
           |||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||||
Db     245 CCAGTTAATTAACCCATTAGGACTTGTGCTGTTCTTGTTTTGTTTCTTTTGCGTGAAGGG 304

Qy     822 CAAGGACACCATTATT 837
           ||||||||||||||||
Db     305 CAAGGACACCATTATT 320
```

```
RESULT  11
BF018105/c
LOCUS       BF018105      458 bp    mRNA            EST       10-OCT-2000
DEFINITION  uw95h11.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:3469797 3', mRNA sequence.
ACCESSION   BF018105
VERSION     BF018105.1  GI:10749437
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uw95h11.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1389157
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3469797"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      123 a    106 c     85 g    144 t
ORIGIN
```

```
  Query Match             24.5%;  Score 295;  DB 143;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 9.7e-75;
  Matches  295;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     909 TCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGT 968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     458 TCGAGCCCTTGGCCTTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGT 399
```

```
Qy     969 GGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCC 1028
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     398 GGCATCGAGAAACTGTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCC 339

Qy    1029 ATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAA 1088
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     338 ATTATCTGTAAGAAAAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAA 279

Qy    1089 ATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGT 1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 ATGTGCAGATGCACAGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGT 219

Qy    1149 TCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 1203
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     218 TCTAACAGAATCTAGTCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 164


RESULT 12
AA418955
LOCUS       AA418955      492 bp    mRNA            EST       16-OCT-1997
DEFINITION  zw01c10.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:768018
            5', mRNA sequence.
ACCESSION   AA418955
VERSION     AA418955.1  GI:2080765
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:768018"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      117 a    136 c    136 g    103 t
ORIGIN


  Query Match           24.2%;  Score 290.8;  DB 7;  Length 492;
  Best Local Similarity   84.1%;  Pred. No. 1.7e-73;
  Matches  328;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;
```

```
Qy     281 CTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAATGTGCCCCGTATCCAGTGTGAAGAT 340
           |||||||||| |  ||||||||||||| |||| |||| |||| |||||||||| ||||
Db     103 CTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCCATATCCAGTGTGGAGAT 162

Qy     341 GGTTGTGACCCACAAGGACTCAAGGACAACAGCCAGTTCTGCTTGCAAAGGATCCGCCAA 400
           || |||||||| |||||||||| ||||||||| |||||||||||||||||||||| |||
Db     163 GGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGCAAAGGATCCACCAG 222

Qy     401 GGTCTGGCTTTTTATAAGCACCTGCTTGACTCTGACATCTTCAAAGGGGAGCCTGCTCTA 460
           ||||||  ||||||| || | ||||| |  || || || |||| |||||||||| ||||
Db     223 GGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGGAGCCTTCTCTG 282

Qy     461 CTCCCTGATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCAGCCAACTCCTC 520
           ||||||||||||||  ||| ||| ||||| |||||||||| || ||||||||||||||
Db     283 CTCCCTGATAGCCCTGTGGCGCAGCTTCATGCCTCCCTACTGGGCCTCAGCCAACTCCTG 342

Qy     521 CAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTTCTAGTCAGCAG 580
           ||||| |||| ||||| | ||||||| || ||||| || ||||| ||| | || ||||
Db     343 CAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCCAGCCA 402

Qy     581 TGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCTTTTTGGCCATA 640
           |||||||| | ||||||||| | ||| |||||||| |||||||||||||| |||| ||
Db     403 TGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTGTA 462

Qy     641 GCTGCCCGGGTCTTTGCCCACGGAGCAGCA 670
           || |||||||||||||||| |||||||||
Db     463 GCCGCCCGGGTCTTTGCCCATGGAGCAGCA 492


RESULT 13
AW987117/c
LOCUS       AW987117      447 bp    mRNA            EST       02-JUN-2000
DEFINITION  uf14d07.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1511341 3', mRNA sequence.
ACCESSION   AW987117
VERSION     AW987117.1  GI:8181925
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:938193.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1511341"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      120 a    102 c     79 g    146 t
ORIGIN
```

```
  Query Match           23.4%;  Score 281;  DB 123;  Length 447;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-70;
  Matches  281;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     923 TTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACT 982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447 TTGGAAGCGGGCAAGCAGCTGCGTGGCCTGAGGGGAAGGGGGCGGTGGCATCGAGAAACT 388

Qy     983 GTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAA 1042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387 GTGAGAAAACCCAGAGCATCAGAAAAAGTGAGCCCAGGCTTTGGCCATTATCTGTAAGAA 328

Qy    1043 AAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCAC 1102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 AAACAAGAAAAGGGGAACATTATACTTTCCTGGGTGGCTCAGGGAAATGTGCAGATGCAC 268

Qy    1103 AGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTA 1162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 AGTACTCCAGACAGCAGCTCTGTACCTGCCTGCTCTGTCCCTCAGTTCTAACAGAATCTA 208

Qy    1163 GTCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 1203
           |||||||||||||||||||||||||||||||||||||||||
Db     207 GTCACTAAGAACTAACAGGACTACCAATACGAACTGACAAA 167


RESULT  14
C06368
LOCUS       C06368        372 bp    mRNA            EST       16-OCT-1996
DEFINITION  C06368 Human pancreatic islet Homo sapiens cDNA clone hbc5699, mRNA
            sequence.
ACCESSION   C06368
VERSION     C06368.1  GI:1503144
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   Takeda,J.
  TITLE     Human pancreatic islet ESTs
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Jun Takeda
            Institute for Molecular and Cellular Regulation, Gunma University
            3-39-15 Showa-machi, Maebashi Gunma 371, Japan
            Tel: 272-20-8856
            Fax: 272-20-8896
            Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hbc5699"
                     /clone_lib="Human pancreatic islet"
                     /note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
                     I; mRNA was prepared from normal adult human islets. cDNA
                     was directionally synthesized from the Xho I in the vector
                     to the EcoRI site. cDNA was size fractionated to remove
                     sequences <1000 bp in size."
BASE COUNT      107 a     88 c    116 g     59 t      2 others
ORIGIN


  Query Match           15.3%;  Score 183.8;  DB 155;  Length 372;
  Best Local Similarity  77.3%;  Pred. No. 2.2e-42;
  Matches  248;  Conservative    0;  Mismatches   69;  Indels    4;  Gaps    2;

Qy      18 ACAGAGTTAGACTCAGAACCAAAGGAGGTGGATAGGGGGTCCA-CAGGCCTGGTGCAGAT 76
           | ||| | || |||| | | |  ||| ||||  || |  |||| |||| ||||||||
Db      55 AGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTGGTGCAAGG 114

Qy      77 CACAGAGCCAGCCAGATCTGAGAAGCAGGGAACAAGATGCTGGATTGCAGAGCAGTAATA 136
           | ||||||||||||||| ||||||| ||| || ||||||||||   ||||||| |||||
Db     115 CGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAGATGCTGGGGAGCAGAGCTGTAATG 174

Qy     137 ATGCTATGGCTGTTGCCCTGGGTCACTCAGGGCCTGGCTGTGCCTAGGAGTAGCAGTCCT 196
           |||| | |||| ||||||||    ||||||||  ||||||||| || | ||||| |||
Db     175 CTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCCCT 234

Qy     197 GACTGGGCTCAGTGCCAGCAGCTCTCTCGGAATCTCTGCATGCTAGCCTGGAACGCACAT 256
           | |||| |||||||||||||||| || | ||| |||||||  || |||||||  ||||||
Db     235 GCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACAT 294

Qy     257 GCACCAGCGGGACATATGAATCTACTAAGAGAAGAAGAGGATGAAGAGACTAAAAATAAT 316
           ||| || |||||| ||| ||   |||||||||||| |  |||||||||||||  ||| ||
Db     295 CCACTAGTGGGACACATGGAT---CTAAGAGAAGANGGAGATGAAGAGACTACNAATGAT 351

Qy     317 GTGCCCCGTATCCAGTGTGAA 337
           || |||| |||||||||||| |
Db     352 GTTCCCCATATCCAGTGTGGA 372


RESULT  15
AI796983/c
LOCUS       AI796983      542 bp    mRNA            EST       18-DEC-1999
DEFINITION  we25e07.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2342148 3',
            mRNA sequence.
ACCESSION   AI796983
VERSION     AI796983.1  GI:5362455
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 868   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2342148"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      150 a    125 c    140 g    127 t
ORIGIN


  Query Match           14.3%;  Score 172.4;  DB 102;  Length 542;
  Best Local Similarity  75.1%;  Pred. No. 5.2e-39;
  Matches  257;  Conservative    0;  Mismatches   76;  Indels    9;  Gaps    3;
```

```
Qy    450 AGCCTGCTCTACTCCCTGATAGCCCCATGGAGCAACTTCACACCTCCCTACTAGGACTCA 509
          ||||| |||| ||||||||||||||||  |||  || ||||.  |||||||||| || ||||
Db    542 AGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCA 483

Qy    510 GCCAACTCCTCCAGCCAGAGGATCACCCCCGGGAGACCCAACAGATGCCCAGCCTGAGTT 569
          |||||||||| ||||| |||| ||||| | ||||||| || ||||| || ||||| |||
Db    482 GCCAACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTC 423

Qy    570 CTAGTCAGCAGTGGCAGCGCCCCCTTCTCCGTTCCAAGATCCTTCGAAGCCTCCAGGCCT 629
          | || ||||  |||||||| | ||||||||| | ||| |||||||| ||||||||||||
Db    422 CCAGCCAGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCT 363

Qy    630 TTTTGGCCATAGCTGCCCGGGTCTTTGCCCACGGAGCAGCAACTCTGACTGAGCCCTTAG 689
          || ||||  |||| |||||||||||||||||| ||||||||||     ||||| || ||
Db    362 TTGTGGCTGTAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAAC----CCTGAGTCCCTAA 307

Qy    690 TGCCAACAGCTTAAGGATG---CCCAGGTTCCCATGG--CTACCATGATAAGACTAATCT 744
           | || ||||| |||||||   | ||| |  ||||||  |  | | |  |||| |||||
Db    306 AGGCAGCAGCTCAAGGATGGCACTCAGATCTCCATGGCCCAGCAAGGCCAAGATAAATCT 247

Qy    745 ATCAGCCCAGACATCTACCAGTTAATTAACCCATTAGGACTT 786
          | || ||||| || ||   ||  ||       |||||  | |
Db    246 ACCACCCCAGGCACCTGTGAGCCAACAGGTTAATTAGTCCAT 205
```

Search completed: June 21, 2001, 01:25:02
Job time: 7692 sec